Diplôme national de master

Domaine - sciences humaines et sociales

Mention – sciences de l'information et des bibliothèques

Parcours – archives numériques

Mémoire / septembre 2017

L'accès aux archives publiques dans le monde numérique

Xavier Capitaine

Sous la direction de Céline Guyon

Remerciements

Je tiens d'abord à adresser mes remerciements à ma directrice de mémoire, Céline Guyon, avec qui cette recherche s'est progressivement construite et affinée tout au long de l'année. Je vous suis reconnaissant pour vos conseils et votre soutien.

Merci à mes relecteurs pour le temps passé à me lire et pour leurs corrections et remarques qui m'ont permises d'améliorer mon mémoire.

Enfin, j'adresse mes remerciements à tous les archivistes que j'ai contactés et qui ont accepté de répondre à mes questions. Leurs réponses m'ont fourni une aide précieuse.

Droits d'auteur réservés. **OU** (cc) BY-NC-ND

Résumé :

Ce mémoire s'intitule : l'accès aux archives publiques dans le monde numérique. Dans ce mémoire nous associons archives publiques et accès afin de constater l'impact du numérique dans ce domaine. La thématique de l'accès renvoie à plusieurs notions bien différentes : c'est à la fois l'accès physique, de la recherche via un portail ou un site à l'exploitation et la réutilisation, mais aussi l'accès règlementaire : la diffusion en ligne est très encadrée. Notre problématique est la suivante : comment se traduit l'évolution des formes d'accès aux archives publiques dans l'environnement numérique ? Ce mémoire a pour but de donner une vision étendue de la thématique de l'accès aux archives publiques numériques et tentera de répondre à de nombreuses interrogations.

Abstract :

This thesis is entitled: access to public archives in the digital world. In this thesis we associate public archives and access in order to see the impact of digital. The theme of access refers to several notions: it is both physical access, from research through a portal or website to exploitation and re-use, but also regulatory access: dissemination online is very framed. Our topic is the following: how evolved the access forms to public archives in the digital environment? The purpose of this thesis is to present an extended view to the issue of access to digital public archives and to try to answer many questions.

Descripteurs : Accès, diffusion, communication, archives publiques, open data, règlementation, droit à l'oubli, déréférencement.

Keywords : access, dissemination of information, communication, public archives, open data, regulation, right to forget, dereferencing.

Droits d'auteur réservés. OU (cc) BY-NC-ND

Droits d'auteurs

Droits d'auteur réservés.

Toute reproduction sans accord exprès de l'auteur à des fins autres que strictement personnelles est prohibée.

OU

Cette création est mise à disposition selon le Contrat : « **Paternité-Pas d'Utilisation Commerciale-Pas de Modification 4.0 France** » disponible en ligne http://creativecommons.org/licenses/by-nc-nd/4.0/deed.frou par courrier postal à Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.

Sommaire

SIGLES ET ABRÉVIATIONS 9

INTRODUCTION..... 11

CHAPITRE I : LA MISSION TRADITIONNELLE DE COMMUNICATION DES ARCHIVES PUBLIQUES 15

I> historique de l'accès aux archives 15

A> La loi du 7 messidor an II..... 15

B> Du XIXème siècle au 3 janvier 1979 16

II> La règlementation du droit d'accès 18

A> Les délais de communicabilité 18

B> Les accès particuliers..... 22

C> Le rôle de la CADA 26

III> les modalités d'accès matériel aux archives publiques, dans le contexte papier..... 31

A> Les dispositions législatives..... 32

B> La salle de lecture 32

C> L'aide à la recherche 33

D> Les chiffres de 2015..... 34

CHAPITRE II : LA RÉVOLUTION NUMÉRIQUE OU LE PASSAGE DE LA COMMUNICATION À LA DIFFUSION..... 37

I> Un nouveau contexte – tant législatif que sociétal – favorable à l'ouverture des données publiques..... 37

A> L'ouverture des données..... 38

B> Évolutions législatives et règlementaires 44

II> Les nouvelles formes d'accès..... 52

A> Les portails open data 52

B> Les services publics d'archives et le numérique 60

III> Vers de nouveaux acteurs ? 74

A> Le cas d'un EPIC : la SNCF 75

B> Le cas d'une entreprise privée : Filae..... 81

CHAPITRE III : UN ACCÈS TRANSFORMÉ..... 85

I> L'accès autour des données à caractère personnel 85

A> L'autorisation unique 029 et la règlementation européenne 86

B> Le déréférencement 89

C> La tension entre mémoire et droit à l'oubli 93

II> De nouvelles pratiques d'accès..... 96

Droits d'auteur réservés. OU (cc) BY-NC-ND

A> L'impact du numérique sur la méthodologie de recherche archivistique 96

B> La réutilisation.....100

C> Des services innovants : le cas des AD des Hautes-Alpes 103

CONCLUSION 111

SOURCES.....115

BIBLIOGRAPHIE.....117

ANNEXES.....121

TABLE DES MATIÈRES.....139

Droits d'auteur réservés. OU (cc) BY-NC-ND

Sigles et abréviations

AD	Archives départementales
AM	Archives municipales
AN	Archives nationales
API	*Application programming interface*
CADA	Commission d'accès aux documents administratifs
CJUE	Cour de justice de l'Union européenne
CNIL	Commission nationale de l'informatique et des libertés
CRPA	Code des relations entre le public et l'administration
EPIC	Établissement public à caractère industriel et commercial
IR	Instrument de recherche
LO	Licence ouverte
MAJ	Mise à jour
ODbL	*Open Database License*
SIAF	Service interministériel des Archives de France

Droits d'auteur réservés. OU (cc) BY-NC-ND

INTRODUCTION

« L'accès aux archives est un des fondements de l'égalité des droits et de l'accès aux savoirs pour les citoyens et constitue donc un des principes de notre République[1] ». C'est ainsi qu'Audrey Azoulay – alors Ministre de la Culture et de la Communication – signifiait l'importance de l'accès aux archives publiques en 2014.

Avant de commencer à débattre autour de cette thématique, comment pouvons-nous définir l'accès ? La définition du Larousse nous dit qu'il s'agit de la « possibilité pour quelqu'un, pour un groupe, d'accéder à une connaissance, de la posséder et de la maîtriser[2] ». L'accès est donc indissociable de la connaissance.

Les services publics d'archive sont un acteur du partage de ces connaissances. En effet, classiquement la communication est le 4ème « C » des missions des archivistes (après la collecte, le classement et la conservation). La grande mission de conservation des archives a une fin en soi : la communication. Celle-ci justifie toutes les autres missions de la profession. Il n'est utile de conserver que dans la mesure où des citoyens ont ensuite la possibilité de consulter. L'abrégé d'archivistique nous dit de la communication qu'elle « consiste à mettre les archives, en fonction de leur communicabilité et de leur état matériel de conservation, à disposition du public et des administrations, soit sur place dans la salle de lecture, soit à distance[3] ». Cette définition répond bien à celle de l'accès. Rappelons que selon le code du patrimoine, les archives sont « l'ensemble des documents, y compris les données, quels que soient leur date, leur lieu de conservation, leur forme et leur support, produits ou reçus par toute personne physique ou morale et par tout service ou organisme public ou privé dans l'exercice de leur activité[4] ».

Dans ce mémoire, nous traiterons exclusivement des archives publiques. Celles-ci rassemblent les documents issus de l'activité de l'Etat, des collectivités territoriales,

---

[1] https://francearchives.fr/file/b0d6555950508ab637adb10ece33d381644d6d37/2017_03_24_RAPPORT_DEFINITIF_NOUGARET.compressed.pdf

[2] Dictionnaire en ligne Larousse : http://www.larousse.fr/dictionnaires/francais/acc%C3%A8s/420

[3] CHAVE Isabelle (sous la direction de) et collectif d'auteurs, *Abrégé d'archivistique, principes et pratiques du métier d'archiviste*, Paris, AAF, 2012.

[4] Article L211-1 du code du patrimoine

Droits d'auteur réservés. OU (cc) BY-NC-ND

des établissements publics, des autres personnes morales de droit public ou des personnes de droit privé chargées de la gestion d'un service public, dans le cadre de leur mission de service public[1].

Afin d'ancrer ce mémoire dans la réalité, nous avons réalisé une enquête auprès d'archivistes travaillant dans différents services publics d'archives autour de problématiques liées à l'accès. Nous avons ainsi interrogé des personnes provenant des Archives municipales, des Archives départementales, des Archives de France et de l'équipe du programme Vitam. Ainsi, la diversité des services interrogés – qui exercent des missions bien différentes autour des archives – nous permet de porter un regard global sur cette thématique et de constater s'il existe, ou non, un regard différent à l'accès selon les services. Des éléments de réponse apportés par ces archivistes seront utilisés tout au long de ce mémoire. Les questionnaires complétés sont disponibles en annexe de ce présent mémoire.

A la question « quelle est votre définition de l'accès aux archives publiques », tous ces archivistes ont insisté sur le fait que l'accès repose d'abord sur un droit du citoyen, dans l'exercice de la démocratie.

Par exemple pour Cyril Longin, directeur des AM de Saint Etienne, « le citoyen a librement et gratuitement accès aux archives. C'est la notion fondamentale de l'accès aux archives ».

Pour Marie Ranquet, conservateur du patrimoine et adjointe au chef du bureau du contrôle et de la collecte des archives publiques aux Archives de France, l'accès est « avant tout une question de démocratie et d'exercice des droits. Le côté « ludique » ou « recherche », même scientifique, passe au second plan ».

Pour Jeanne Mallet, conservateur du patrimoine et chargée de l'accès aux archives au SIAF, « il s'agit de permettre à toute personne qui en justifie le besoin, ou simplement l'envie, la consultation de documents utiles à son histoire personnelle, à ses recherches scientifiques ou qui lui sont nécessaires pour prouver un droit. Ce dernier aspect est d'ailleurs, historiquement et juridiquement, la justification première du droit d'accès aux archives tel qu'il est conçu en France. C'est un rouage essentiel de notre démocratie ».

[1] Article L211-4 du code du patrimoine

Droits d'auteur réservés. OU (cc) BY-NC-ND

Enfin, pour Mélanie Rebours, conservateur du patrimoine et directrice de la diffusion et des partenariats du Programme VITAM, et pour son équipe, l'accès aux archives publiques correspond à la « capacité des citoyens et autres utilisateurs à rechercher et consulter l'information produite par les administrations ou leurs délégataires dans l'exercice de leurs missions de service public, dans le respect du droit d'en connaître. »

Cela nous donne un premier aperçu de la manière dont est perçu l'accès aux archives publiques. Mais pour ce mémoire, nous avons décidé de nous intéresser à l'accès aux archives publiques dans un cadre bien spécifique, celui du numérique.

En effet, depuis le début du XXIème siècle et le développement d'internet, nous avons pu constater des changements, l'accès change de forme. Dans un environnement papier, l'accès au document peut être difficile. Il nécessite un déplacement, du temps, et il peut être restreint pour différentes raisons comme par exemple dans un but de préservation. Mais le numérique a permis de surpasser cette problématique par le biais de la numérisation. Certaines barrières de l'accès sont ainsi supprimées avec internet : désormais un document peut être accessible directement de chez soi, et même consulté de manière simultanée par plusieurs personnes. Le modèle de la consultation en salle de lecture est désormais détrôné par celui de l'accès en ligne.

Autre transformation, dans l'univers papier le lecteur se déplace dans un service public d'archives afin d'obtenir des réponses à un besoin particulier : il peut s'agir d'une recherche juridique, historique, généalogique etc. Avec le numérique, et notamment un mouvement appelé open data, la démarche est inversée. L'administration va diffuser l'information d'une manière proactive, sans répondre à une demande précise. Pour cela elle anticipe les besoins des internautes pour propager la connaissance auprès d'un public élargi.

De plus, un nouveau vocabulaire est apparu avec le numérique. On parle d'informations publiques, ou de données publiques. Selon l'administrateur général des données – un poste également créé par le numérique – « une donnée est la description élémentaire de nature numérique, représentée sous forme codée, d'une réalité en vue d'être collectée, enregistrée, traitée, manipulée, transformée,

Droits d'auteur réservés. OU (cc) BY-NC-ND

conservée, archivée, échangée, diffusée, communiquée[1] ». Attention, les données publiques ne correspondent pas à l'ensemble des données accessibles en ligne. En effet, les données publiques « concernent les informations ou données produites ou reçues par une autorité administrative dans le cadre de sa mission de service public, publiées par une autorité administrative ou communicables à toute personne en faisant la demande. Ces informations doivent être présentées sous un format permettant leur traitement automatisé et leur réutilisation [2]». De même, le terme de lecteur disparait pour laisser place à ceux d'internaute, d'utilisateur ou de réutilisateur. La « diffusion » tend également à remplacer le terme de « communication ». Le mot « ouverture » est aussi souvent utilisé.

Notre problématique est donc la suivante : comment se traduit l'évolution des formes d'accès aux archives publiques dans l'environnement numérique ?

Nous allons tenter de répondre à de nombreuses interrogations. Pourquoi donner accès en ligne à des archives publiques ? Comment la législation a-t-elle évoluée suite à la révolution numérique et sous quelles conditions permet-elle l'accès aux archives publiques ? À quelles informations publiques le citoyen a-t-il accès, et par quels médias ? Qui sont les acteurs de cette diffusion ? Comment répondre aux demandes du public virtuel ? Quel accès mettre en place dans un contexte de tension entre mémoire et droit à l'oubli ? Peut-on à la fois donner accès et garantir la protection de la vie privée ?

Dans une première partie nous nous intéresserons à la mission traditionnelle de communication exercée par les services publics d'archives. Puis, nous traiterons du passage de la communication à la diffusion dans un deuxième chapitre. Enfin, le dernier chapitre sera consacré au bouleversement du rapport à l'accès apporté par le numérique.

[1] https://agd.data.gouv.fr/

[2] Vade-mecum sur l'ouverture et le partage des données publiques : http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/vademecum-ouverture.pdf

Droits d'auteur réservés. OU (cc) BY-NC-ND

CHAPITRE I : LA MISSION TRADITIONNELLE DE COMMUNICATION DES ARCHIVES PUBLIQUES

Comme nous l'avons signalé en introduction, la communication est le 4ème « C » des missions des archivistes. Les archivistes travaillent pour des lecteurs, qui attendent une restitution de l'information. Si on se place du côté du service d'archive, on parle de communication. Du point de vue des utilisateurs, on parle plutôt d'accès aux archives ou de consultation.

Après un rapide retour historique sur l'accès aux archives publiques (I), nous nous attacherons à dresser un panorama le plus exhaustif possible de la règlementation du droit d'accès (II), pour enfin clore ce chapitre par les modalités concrètes d'un accès matériel aux archives publiques (III).

I> HISTORIQUE DE L'ACCÈS AUX ARCHIVES

C'est de la révolution française que nait la notion d'archive publique. En effet c'est lors de cette période qu'est créée une administration spécifique des archives pour répondre à de nouveaux besoins tels que la conservation des archives des institutions nouvelles, mais aussi la prise en charge des archives de l'Ancien régime saisies comme biens nationaux. Afin de protéger cette nouvelle administration, une règlementation est rapidement instaurée (A), celle-ci ne faisant qu'évoluer jusqu'à la fameuse loi sur les archives du 3 janvier 1979 (B).

A> La loi du 7 messidor an II

Un des textes fondamentaux est la loi du 7 messidor an II (25 juin 1794). Ce texte fondateur s'organise autour de quelques idées principales comme le principe de centralisation des archives de la Nation ou encore celui de la constitution d'un outil d'élaboration de la mémoire nationale. Mais ce n'est pas tout, un des grands mérites attribué à cette loi de messidor est d'avoir instauré la publicité des archives, objet de l'article 37, l'un des plus connus de toute la loi[1] :

> XXXVII. Tout citoyen pourra demander dans tous les dépôts, aux jours et heures qui seront fixés, communication des pièces

[1] Légilux : http://legilux.public.lu/eli/etat/leg/loi/1794/06/25/n1/jo

Droits d'auteur réservés. OU (cc) BY-NC-ND

> qu'ils renferment : elle leur sera donnée sans frais et sans déplacement, et avec les précautions convenables de surveillance. Les expéditions ou extraits qui en seront demandés seront délivrés à raison de quinze sous du rôle.

C'est le passage des archives d'État à celles de la Nation. Ces archives doivent être mises à la disposition des citoyens. Le principe de publicité se substitue donc à celui du secret d'Etat qui les caractérisait auparavant. Ainsi, le but est de montrer la transparence des décisions de l'Assemblée constituante.

Cette loi est une manière de donner des armes à l'article XV de la Déclaration des Droits de l'Homme et du Citoyen[1] :

> La société a le droit de demander compte à tout agent public de son administration.

Pour certains, cette fameuse « publicité » des archives marque le véritable apport de la Révolution française dans le domaine de l'archivistique[2]. La Constituante ou la Convention n'autorisèrent en fait qu'un accès extrêmement limité à quelques catégories de documents seulement. En effet, la consultation aux Archives nationales est restée soumise durant tout l'Empire à même au-delà à l'autorisation du ministre de l'Intérieur.

B> Du XIXème siècle au 3 janvier 1979

Le premier tiers du XIXème siècle est marqué par une stagnation, voire une régression. Ainsi, le principe de la libre communicabilité des documents d'archives est de plus en plus réduit par des règlements internes[3].

Ce n'est que vers le milieu du XIXème siècle que les Archives nationales ne se comportent d'une salle de consultation, quand la recherche dans les archives est devenue une nécessité de la science historique.

[1] Légifrance : https://www.legifrance.gouv.fr/Droit-francais/Constitution/Declaration-des-Droits-de-l-Homme-et-du-Citoyen-de-1789

[2] Pierre Santoni, archives et violence. A propos de la loi du 7 messidor an II, dans la Gazette des archives, 1989, volume 146, p199-214.

[3] CHAVE Isabelle (sous la direction de) et collectif d'auteurs, *Abrégé d'archivistique, principes et pratiques du métier d'archiviste*, Paris, AAF, 2012

Droits d'auteur réservés. OU (cc) BY-NC-ND

A la fin du XIXème siècle, plusieurs lois prévoient des mesures de publicité, notamment par affichage, pour des décisions administratives[1].

Mais il faut attendre 1960 pour voir surgir une véritable revendication citoyenne pour disposer d'un droit d'accès aux documents administratifs, où trois arguments sont notamment avancés[2] :

- Le public, qui dispose d'une meilleure éducation, revendique « le droit de savoir ».
- Ce serait pour l'administration, qui est critiquée pour son goût du secret, un moyen d'améliorer son image.
- Enfin, mettre à disposition l'information à tous les citoyens serait un moyen pour l'administration de faire adhérer la population aux projets collectifs.

C'est à la fin des années 1970 que le législateur prend acte de ces revendications et souhaite renforcer la transparence administrative, en publiant notamment trois grandes lois sur l'accès aux archives publiques.

Tout d'abord, est votée la loi n° 78-17 du 6 Janvier 1978 relative à l'informatique, aux fichiers et aux libertés. Cette loi établit pour les personnes physiques un droit d'accès et de rectification aux fichiers nominatifs qui les concernent.

Ensuite, intervient la loi n° 78-753 du 17 juillet 1978 qui pose le principe du libre accès aux documents administratifs.

Enfin, la loi n° 79-18 du 3 janvier 1979 sur les archives fixe les délais aux termes desquels les archives deviennent librement consultables. Elle consacre ainsi le « droit de savoir » et le « droit à la mémoire ». Cette loi témoigne d'une volonté d'ouverture des archives et de libéralisation des délais de communicabilité.

En définitive, la loi du 7 messidor a duré près de deux siècles, n'ayant été remplacée que par la loi du 3 janvier 1979. Elle est donc la base de tout ce qui a été créé en France relativement aux archives. Jean Favier écrit d'ailleurs que, pour les archives françaises,

[1] CHAVE Isabelle (sous la direction de) et collectif d'auteurs, *Abrégé d'archivistique, principes et pratiques du métier d'archiviste*, Paris, AAF, 2012

[2] Site Web de la CADA : http://www.cada.fr/historique-de-loi-du-17-juillet-1978,6054.html

Droits d'auteur réservés. OU (cc) BY-NC-ND

« le décret de messidor an II était resté et allait demeurer jusqu'en 1979 la charte fondamentale[1] ».

II> LA RÈGLEMENTATION DU DROIT D'ACCÈS

Un document d'archive conservé a vocation à être diffusé. Mais il n'est pas toujours immédiatement communicable selon son contenu ou sa nature. C'est pourquoi depuis le 3 janvier 1979 de nombreux autres textes de loi ont été votés pour instaurer des délais de communicabilité, afin de protéger à la fois les intérêts de l'État mais aussi les droits et libertés individuels. Certains documents ne sont donc accessibles au public qu'au terme de ces délais fixés par la loi (A).

Cependant, il existe des exceptions à cette règle, de façon à autoriser un accès anticipé, ou extraordinaire, selon les documents pour un objet de recherche, selon la profession de ceux qui en réclament l'accès ou encore selon l'identité même de ces personnes et leur lien direct avec l'archive (B).

Une instance est chargée de veiller à la transparence de l'action administrative et au libre accès de tout citoyen aux documents administratifs, il s'agit de la CADA (C).

A> Les délais de communicabilité

À la fin des années 1990, un constat sévère est porté sur la communicabilité des archives, constat issu du rapport Braibant. En effet, en 1995, Guy Braibant[2] est chargé de dresser le « bilan de l'application de loi[3] » de 1979, en particulier en ce qui concerne les restrictions d'accès auxquelles sont soumises certaines catégories de documents. Le rapport, intitulé *Les archives en France*, est remis à Alain Juppé l'année suivante. Dans ce rapport, Guy Braibant réclame une réforme en profondeur, car selon lui « la règle législative très libérale » de 1979 est en réalité « vidée de sa portée par des textes réglementaires ». La conséquence directe qu'il dénonce est une restriction d'accès aux archives publiques.

Quelques années après ce rapport[4], le gouvernement décide finalement de libéraliser le droit des archives. Aujourd'hui, l'accès aux archives publiques et documents

[1] FAVIER Jean, *les archives*, PUF, que sais-je ?, 2001

[2] Guy Braibant est alors président de section honoraire au Conseil d'Etat.

[3] COEURE Sophie, DUCLERT Vincent, *Les archives*, Paris, La découverte, 2011.

[4] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

administratifs est fixé à travers deux codes, dont il sera fait souvent mention dans ce mémoire. Il s'agit du Code du patrimoine (2004), et du Code des relations entre le public et l'administration (2015).

En ce qui concerne le Code du patrimoine, le régime de communication des archives publiques est fixé par le chapitre 3 du livre II du Code du patrimoine dans les articles L213-1 à L213-8.

Quant au Code des relations entre le public et l'administration, c'est le livre III intitulé « l'accès aux documents administratifs et la réutilisation des informations publiques », qui précise le droit de toute personne à l'information. La liberté d'accès aux documents administratifs y est précisée et garantie par les dispositions des titres Ier, III et IV[1] :

- Titre Ier : le droit d'accès aux documents administratifs
- Titre III : Les personnes responsables de l'accès aux documents administratifs et des questions relatives à la réutilisation des informations publiques
- Titre IV : La commission d'accès aux documents administratifs

En plus des grandes règles établies par ces deux textes, qui sont liés de manière étroite pour former un cadre général cohérent, il faut noter qu'il existe pléthore d'autres lois qui créent des dispositions particulières d'accès à certains types d'informations ou de documents : il est donc nécessaire de toujours vérifier si un droit particulier peut s'appliquer à tel ou tel document. Nous pouvons retrouver ces lois qui contiennent des dispositions particulières par exemple dans le Code de l'environnement, le Code électoral, le Code de procédure fiscale, le Code de procédure pénal, le Code de la santé publique…

Le livre II du Code du patrimoine a été modifié en profondeur par la loi n°2008-696 du 15 juillet 2008 dont le but était d'homogénéiser les dispositions contenues dans la loi n°79-18 du 3 janvier 1979 sur les archives et celles de la loi n°78-753 du 17 juillet 1978 portant diverses mesures d'amélioration des relations entre l'administration et le public et diverses dispositions d'ordre administratif, social et fiscal, dite « loi CADA », et codifiée depuis 2016 dans le Code des relations entre le public et l'administration. L'impératif du

[1] Légifrance : https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000031366350

Droits d'auteur réservés. OU (cc) BY-NC-ND

projet, selon les mots de Christine Albanel, alors ministre la culture et la communication, était de « faciliter et d'accélérer l'accès de tous aux archives publiques »[1].

Les délais de communicabilité votés à ce moment-là, en dernière lecture, établissent un régime plus libéral en matière de communication. Le délai de 30 ans qui prévalait jusque-là est supprimé. En effet avant 2008 l'article L213-1 du Code du patrimoine indiquait[2] :

> […] Tous les autres documents d'archives publiques pourront être librement consultés à l'expiration d'un délai de trente ans.

Depuis 2008, l'article L213-1 est désormais le suivant :

> Les archives publiques sont, sous réserve des dispositions de l'article L. 213-2, communicables de plein droit.

Cette loi de 2008 est une victoire pour Christine Albanel, pour qui « désormais, chaque Français pourra consulter librement et immédiatement les archives publiques[3] ». Selon elle cette disposition renoue avec « les principes affirmés à la Révolution, établissant pour chaque citoyen d'avoir accès gratuitement et librement aux documents produits par l'administration et de contrôler par là même l'action de cette dernière[4] ».

Elle clos son discours en rappelant qu'« il aura fallu plus de deux siècles pour que soit rétabli ce principe fondamental, gage d'une réelle démocratie »[5].

Le principe retenu est donc le suivant : les archives publiques sont communicables à tous et sans délai, mais il existe des exceptions.

L'article L213-2 du Code du patrimoine énumère ces exceptions à la règle de libre communicabilité et fixe des délais de communicabilité. Ceux-ci sont plus ou moins longs, en fonction du contenu du document, et vise à les protéger. Ces délais vont de 25 à 120 ans. Par exemple, est prévu un délai de 50 ans pour les documents contenants une « appréciation ou un jugement de valeur sur une personne physique, nommément

[1] COEURE Sophie, DUCLERT Vincent, *Les archives*, Paris, La découverte, 2011.

[2] Légifrance : https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000006074236

[3] Sophie Coeuré, op.cit.

[4] Ibid.

[5] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

désignée ou aisément identifiable[1] ». Le délai de 120 ans est consacré aux « documents portant atteinte au secret médical[2] » (dans ce cas, le délai est décompté depuis la date de naissance de l'individu, et non depuis la date du document). Tous ces délais sont disponibles sur le portail France Archives[3].

Le Code des relations entre le public et l'administration, dans son article L311-5 fait également mention de documents qui ne sont pas immédiatement communicables[4] :

> Les documents administratifs dont la consultation ou la communication porterait atteinte :
>
> - Au secret des délibérations du Gouvernement et des autorités responsables relevant du pouvoir exécutif
> - Au secret de la défense nationale
> - A la conduite de la politique extérieure de la France
> - A la sûreté de l'État, à la sécurité publique, à la sécurité des personnes ou à la sécurité des systèmes d'information des administrations
> - A la monnaie et au crédit public…

Outre ces secrets d'État, sont également protégés les secrets des personnes (notamment la vie privée et le secret médical), mais aussi le secret industriel et commercial, le secret de l'instruction judiciaire, etc. Une fois ces délais expirés, les documents deviennent accessibles à toute personne qui en fait la demande.

Il faut également rappeler que c'est bien le type d'information qui détermine la communication (ou non) des documents, et non la typologie de ces documents. Il existe seulement quelques exceptions à cette règle, à savoir l'état civil, les minutes et répertoires d'officiers publics et ministériels et quelques autres typologies de documents. Cela implique qu'un dossier, qui comprend donc plusieurs documents, peut contenir des informations qui ne sont pas toutes soumises au même délai. Pour avoir accès à un

[1] CHAVE Isabelle (sous la direction de) et collectif d'auteurs, *Abrégé d'archivistique, principes et pratiques du métier d'archiviste*, Paris, AAF, 2012

[2] Ibid.

[3] Portail France Archives : https://francearchives.fr/fr/article/26287562

[4] Légifrance : https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000031366350

Droits d'auteur réservés. OU (cc) BY-NC-ND

document au sein de ce dossier, il convient alors d'occulter les informations qui ne sont pas communicables.

Par exemple, si un individu souhaite consulter une page d'un registre matricule militaire, il peut le faire suite à un délai de 50 ans à compter de la date de la dernière information protégée par le droit de la vie privée sur cette page. Mais ce délai peut se réduire à 25 ans après le décès de la personne concernée. Pour déterminer la communicabilité d'un document, sa consultation par l'archiviste est souvent une première étape.

Enfin, une autre exception à la règle est mentionnée à l'article L213-2 du Code du patrimoine, à savoir qu'il existe une catégorie de documents qui est tout simplement incommunicable au public, quel que soit la date des documents[1] :

> Ne peuvent être consultées les archives publiques dont la communication est susceptible d'entraîner la diffusion d'informations permettant de concevoir, fabriquer, utiliser ou localiser des armes nucléaires, biologiques, chimiques ou toutes autres armes ayant des effets directs ou indirects de destruction d'un niveau analogue.

Cette disposition de la loi de 2008, qui crée finalement un statut de « non-communicabilité », a été sévèrement critiquée. L'AUSPAN (Association des usagers du service public des Archives nationales) écrit que « « sous le prétexte de la sécurité nationale face à la menace terroriste, le législateur porte abusivement atteinte aux droits des citoyens »[2].

B> Les accès particuliers

1) L'accès privilégié

Rappelons enfin que toute personne directement concernée par un dossier peut y avoir accès, et ceci sans délai. En effet la loi donne le droit d'accéder sans formalité préalable à tous les éléments qui concernent un individu.

Pour cela, si par exemple une personne, qui a travaillé dans une administration, souhaite consulter son dossier une fois arrivée à la retraite, celle-ci doit contacter le service dans

[1] Légifrance : https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000006074236

[2] Sophie Coeuré, op.cit.

Droits d'auteur réservés. OU (cc) BY-NC-ND

lequel elle a travaillé, qui est encore susceptible de conserver le dossier. Dans ce cas elle peut prendre rendez-vous pour le consulter ou en demander une copie. Si le dossier a déjà versé aux archives, alors l'individu peut se présenter directement dans le service d'archives et le consulter après s'être inscrit en salle de lecture ou en demander une copie par correspondance. Cependant, si le dossier comporte des éléments relatifs à d'autres personnes, ces éléments seront masqués avant consultation. Cet accès est protégé par l'article L311-6 du code des relations entre le public et l'administration[1].

On retrouve également des accès privilégiés pour certaines personnes dans d'autres textes, tel que le code de la santé publique. En effet celui-ci, dans son article L.1110-4, autorise des ayants droits à accéder au dossier médical d'une personne décédée dans certaines circonstances[2] :

> Le secret médical ne fait pas obstacle à ce que les informations concernant une personne décédée soient délivrées à ses ayants droit […], dans la mesure où elles leur sont nécessaires pour leur permettre de connaître les causes de la mort, de défendre la mémoire du défunt ou de faire valoir leurs droits, sauf volonté contraire exprimée par la personne avant son décès.

Lorsque ces accès privilégiés sont refusés (exemple d'une personne qui ne fait pas partie de la famille proche), c'est régime général de communicabilité des archives qui s'applique.

2) L'accès professionnel

Dans le cadre de leur activité professionnelle, certaines personnes disposent d'un droit d'accès privilégié aux archives publiques. Cet accès, strictement encadré par des dispositions législatives ou réglementaires, s'applique généralement à des typologies particulières de documents.

Deux professions sont particulièrement liées aux archives publiques, il s'agit des notaires et des généalogistes.

Les notaires, qui se placent en tant que mandataires des intéressés ou de leurs descendants ou ascendants, disposent de ce droit d'accès privilégié à certaines typologies de

[1] Légifrance : https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000031366350

[2] Légifrance : https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000006072665

Droits d'auteur réservés. OU (cc) BY-NC-ND

documents, car ceux-ci sont essentiels aux recherches qu'ils effectuent lors de successions. Il s'agit notamment des registres de l'état civil de moins de soixante-quinze ans, des documents fiscaux de moins de 50 ans, ainsi que des registres matricules du recensement militaire de moins de 120 ans[1].

Les généalogistes professionnels, quant à eux, et uniquement s'ils disposent d'un mandat, peuvent obtenir des facilités d'accès équivalentes. Les Archives de France peuvent leur fournir une autorisation de consultation de l'état civil de moins de soixante-quinze ans, si le tribunal de grande instance de Paris a préalablement donné son accord[2].

Ces autorisations sont valables deux ans, et elles doivent être présentées par les généalogistes avant toute consultation de ces documents.

Au-delà de ces deux professions, un accès peut être nettement plus large pour deux autres catégories de professionnels, le juge d'instruction et les services fiscaux.

En effet, l'article 99-3 du code de procédure pénale indique que le juge d'instruction et l'officier de police judiciaire commis par lui ont un accès privilégié aux documents produits ou reçus par les administrations[3].

Quant aux services fiscaux, c'est l'article L. 92 du code de procédure fiscale qui autorise un droit d'accès très étendu aux actes conservés par les services publics d'archives[4].

3) L'accès par dérogation

L'article L213-3 du Code du patrimoine indique la possibilité d'une consultation d'archives publiques avant l'expiration des délais spéciaux. Cette autorisation peut être de portée individuelle[5] :

> L'autorisation de consultation de documents d'archives publiques avant l'expiration des délais fixés au I de l'article L. 213-2 peut être accordée aux personnes qui en font la demande dans la mesure où l'intérêt qui s'attache à la consultation de ces

[1] Portail France Archives : https://francearchives.fr/fr/article/26287583

[2] Ibid.

[3] Légifrance : https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000006071154

[4] Légifrance : https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000006069583

[5] Légifrance : https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000006074236

Droits d'auteur réservés. OU (cc) BY-NC-ND

> documents ne conduit pas à porter une atteinte excessive aux intérêts que la loi a entendu protéger.

Ou générale :

> L'administration des archives peut également, après accord de l'autorité dont émanent les documents, décider l'ouverture anticipée de fonds ou parties de fonds d'archives publiques.

Par exemple, dans le cas des dérogations générales, la plupart des fonds concernant la Seconde Guerre mondiale ont été progressivement ouverts au public à partir de 1997.

Si un chercheur souhaite consulter un document couvert par le délai de communicabilité fixé par loi, celui-ci a la possibilité de faire appel à une procédure de dérogation afin de demander une autorisation de consultation. Selon la CADA, autour de 90% de ces demandes d'autorisation sont accordées chaque année[1].

Cette procédure est encadrée par la circulaire ministérielle du 29 juillet 2010. Avant d'arriver à la décision finale d'accord ou de refus, la demande passe par un certain nombre d'étapes successives.

Tout d'abord, la demande d'autorisation est à adresser auprès du service d'archives qui conserve le ou les documents souhaités. Le demandeur doit décliner son identité, ses motivations précises, et souscrit un engagement comme quoi il ne divulguera aucune information « susceptible de porter atteinte aux droits et intérêts protégés par la loi[2] ». La reproduction des documents est donc généralement interdite, de façon à protéger la vie privée des personnes ou encore la sécurité de l'État.

Après un premier examen de cette demande et du contenu des documents souhaités, le service d'archives formule un avis, qu'il soit favorable ou non, et transmet le formulaire de demande à l'autorité dont émanent les documents.

Une fois que le service producteur a étudié la demande, il retourne au service d'archives le formulaire rempli et donne son accord ou son refus. Le service d'archives envoie alors la totalité du dossier au Service interministériel des Archives de France, qui est le seul habilité à donner une décision. Cependant, si le service producteur a signifié son refus, le

[1] Site web de la CADA : http://www.cada.fr/l-acces-aux-archives-par-derogation%2c6103.html

[2] Portail France Archives : https://francearchives.fr/fr/circulaire/DGP_SIAF_AACR_2010_10

Droits d'auteur réservés. OU (cc) BY-NC-ND

SIAF ne peut pas passer outre. En revanche, il peut toujours opposer un refus à une demande qui a été préalablement acceptée par le service producteur. Dans tous les cas, tout refus doit être motivé, qu'il provienne du SIAF ou du service producteur.

Si accord il y a, l'autorisation compte pour la consultation de documents expressément identifiés par leur cote.

En cas de refus, l'usager peut saisir dans les deux mois la Commission d'accès aux documents administratifs (CADA). Celle-ci dispose de deux mois suivant la réception de la demande pour émettre un avis. Comme il ne s'agit que d'un avis, il n'est d'aucune manière contraignant pour le service d'archive ou le service producteur. Si le refus est toujours maintenu même avec présentation de l'avis favorable de la CADA, un recours au tribunal administratif peut être engagé par le demandeur. Signalons qu'avant tout contentieux, la saisine de la CADA est obligatoire (article 20 de la loi 78-753)[1].

Le SIAF a mis en place récemment un observatoire national des dérogations[2], qui permet de porter une réflexion globale et d'avoir une vue d'ensemble sur les dérogations individuelles acceptées pour la consultation de documents non librement communicables. C'est également pour le SIAF un moyen de mener une politique homogène pour donner plus de cohérence aux décisions prises.

Cet observatoire nous donne quelques chiffres[3] concernant les dérogations en 2015. Les services d'archives relevant de la tutelle du SIAF ont traité au total 2134 demandes de dérogation, qui représentent 30 206 articles sollicités. 93 % des articles demandés par dérogation ont été accordés, 7 % refusés. 32 recours ont été déposés devant la CADA, qui a rendu 16 avis favorables, 1 avis partiellement favorable et 15 avis défavorables.

C> Le rôle de la CADA

1) Historique et missions

La Commission d'accès aux documents administratifs a été créée par la loi 78-753 du 17 juillet 1978. Il s'agit d'une Autorité administrative indépendante chargée d'assurer la bonne application du droit d'accès. Elle est le premier interlocuteur en la matière, que ce soit pour les citoyens ou pour les administrations.

[1] Site web de la CADA : http://www.cada.fr/l-acces-aux-archives-par-derogation%2c6103.html

[2] Portail France Archives : https://francearchives.fr/article/38293

[3] Portail France Archives : https://francearchives.fr/fr/article/26287583

Droits d'auteur réservés. OU (cc) BY-NC-ND

L'article L340-1 du code des relations entre le public et l'administration précise[1] :

> Elle est chargée de veiller au respect de la liberté d'accès aux documents administratifs et aux archives publiques ainsi qu'à l'application du titre II du présent livre[2] dans les conditions prévues par le présent livre et par le titre Ier du livre II du code du patrimoine.

Comme nous l'avons signalé précédemment en lien avec la procédure de dérogation, la CADA rend des avis qui constituent une voie de recours précontentieuse.

Si une personne se voit refuser l'accès à un document administratif ou n'a pas de réponse dans un délai d'un mois après l'avoir demandé, celle-ci peut saisir la CADA qui donnera un avis sur la communicabilité du document. La CADA a également pour mission de fournir des conseils aux administrations sur la mise en œuvre du droit d'accès ou du droit à réutilisation[3].

Si une autorité administrative a un doute sur le caractère communicable d'un document, celle-ci peut également se tourner vers la CADA afin d'avoir des réponses.

En plus des avis et conseils qu'elle donne, elle peut aussi prononcer des sanctions contre des individus quant à la violation des prescriptions de la loi en matière de réutilisation d'informations publiques.

Son rôle est avant tout de veiller à la transparence de l'action administrative.

2) Le rapport d'activité 2015

Il est intéressant de consulter le rapport annuel de la CADA (le dernier publié étant celui de 2015[4]). Son président, Marc Dandelot, nous délivre plusieurs constats.

[1] Légifrance : https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000031366350

[2] Le titre II porte sur la réutilisation des informations publiques, nous approfondirons ce point dans le troisième chapitre.

[3] Site web de la CADA : http://www.cada.fr/la-cada,3.html

[4] Site web de la CADA : http://www.cada.fr/IMG/pdf/cada_rapport_activite_2015.pdf

Droits d'auteur réservés. OU (cc) BY-NC-ND

a) Une augmentation significative du nombre de saisines en 2015

Tout d'abord, il remarque de la part des citoyens « un besoin croissant d'accès à de nombreux documents administratifs, qu'il s'agisse de documents destinés à un usage personnel ou pour apprécier et contrôler l'action des services publics ».

Lorsque ce besoin n'est pas satisfait, les citoyens peuvent montrer de l'impatience et de la frustration. C'est à ce moment qu'ils font appel aux services de la CADA.

Les demandes portent généralement sur des dossiers volumineux, y compris des bases de données. Ces demandes sont très difficiles à satisfaire pour les administrations, et d'autant plus dans un délai imparti très court. Marc Dandelot signale également un manque de moyens chez les administrations pour répondre efficacement aux administrés qui réclament un droit d'accès. En effet, lorsque la demande porte sur un ensemble volumineux de documents, l'administration peut être « désorientée, voire dissuadée, par l'ampleur de la tâche ». Ce qui pousse certaines administrations à ne pas répondre et attendre la saisine de la CADA, qui délivrera la marche à suivre. Face à une telle attitude, la CADA risque vite d'être surchargée elle-même de demandes et de ne pouvoir y répondre dans les temps. L'administration doit donc jouer pleinement son rôle, la CADA n'intervenant qu'en ultime recours.

Suite à ce premier constat, le président de la CADA nous donne une autre clé d'interprétation concernant la hausse du nombre de saisine de la CADA en 2015 par rapport aux années précédentes. Selon lui, « les règles d'accès aux documents administratifs procèdent aujourd'hui d'un droit singulièrement compliqué ». En effet, comme nous l'avons vu précédemment, la règlementation en matière de droit d'accès est exprimée par une législation que l'on peut qualifier de générale (code du patrimoine, code des relations entre le public et l'administration), mais aussi d'une multitude de législations particulières que l'on trouve dans d'autres codes : code de procédure pénale, code de procédure fiscale… Il peut donc être particulièrement difficile de déterminer la communicabilité d'un document. Une deuxième source de complexité est celle de la diversité des secrets protégés par la loi. Le droit à la vie privé est bien fixé et encadré, mais la portée d'autres droits reste plus vague comme le droit au secret commercial et industriel est plus difficile à mesurer.

Ce constat amène le président de la CADA à déclarer que « le droit d'accès [est devenu] un droit d'initiés, alors qu'il devrait procéder de règles simples et universelles ». L'un des

Droits d'auteur réservés. OU (cc) BY-NC-ND

objectifs de la CADA est donc de parvenir à ce que les règles en matière d'accès soient bien comprises et partagées, de façon à restreindre le nombre de saisines.

Ci-dessous, le tableau des chiffres de l'année 2015 de la CADA[1] qui donne l'ampleur du nombre des affaires traitées :

Nombre d'entrées brut/ Affaires totales traitées	**7 222**
- dont dossiers non instruits	1 404
- dont dossiers instruits	5 818
- *avis*	5 591
- *conseils*	227
Désistements	*693*
Répartition des avis rendus	
- Avis favorable	55,1%
- Sans objet (communiqué ou inexistant)	25%
- Avis défavorable	9,8%
- Avis d'incompétence	5,4%
- Irrecevabilité de la saisine	4,8%
Durée moyenne d'instruction des dossiers	*56*
Demandes de renseignements	
- courriers et courriels	5 800
- téléphone	35 appels/jours
- PRADA	1 604

Légende : PRADA : personnes responsables de l'accès aux documents administratifs (communautés de communes, centres hospitaliers, établissements publics territoriaux, organismes privés chargés d'une mission de service public…)

Un second tableau nous montre l'évolution des avis rendus par la Commission, entre 2011 et 2015[2] :

	2011	2012	2013	2014	2015
Avis favorables	46,5	45,37	57,4	49,1	55,1
Demandes sans objet	33,2	31,4	23	32	25
Avis défavorables	8,4	10,39	9,3	8,6	9,8
Avis d'incompétence	6,4	7,8	4,9	6,3	5,4
Demandes irrecevables	5,5	5,7	5,4	4,0	4,8

[1] Site web de la CADA : http://www.cada.fr/IMG/pdf/cada_rapport_activite_2015.pdf

[2] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

L'intitulé « Demandes sans objet » est à prendre au sens de conseil. Cette diminution du taux traduit à l'inverse une augmentation des affaires complexes nécessitant un avis de la Commission. On remarque que le taux d'avis défavorable reste stable, autour de 10%. Les demandes irrecevables ont tendance à diminuer, ce qui montrerait une meilleure compréhension des services rendus par la CADA.

Enfin, un dernier tableau nous montre l'évolution du taux de réponse des administrations suite à un avis favorable de la CADA, entre 2010 et 2015[1] :

2010	65,9
2011	66,7
2012	63,5
2013	47,15
2014	63,15
2015	42,8

Nous pouvons remarquer que ce taux s'est fortement dégradé en 2015. Cette dégradation peut s'expliquer par le fait que des administrations se retrouvent démunies et sont dans l'impossibilité de donner suite à une demande d'accès. C'est le cas notamment des communes (mises en cause à 31.8% des cas en 2015) qui disposent de trop peu de moyens, surtout lorsqu'elles ont affaire à des demandes volumineuses.

Enfin, notons un dernier chiffre, le délai moyen d'instruction des dossiers en 2015 est de 56 jours contre 50 en 2014[2]. Cette augmentation s'explique par la complexité croissante des affaires.

Seule une exception à la règle existe concernant l'obligation pour une administration de communiquer un document administratif ayant fait la demande d'un particulier, que l'on retrouve dans l'article L311-2 du code des relations entre le public et l'administration[3] :

> L'administration n'est pas tenue de donner suite aux demandes abusives, en particulier par leur nombre ou leur caractère répétitif ou systématique.

[1] Site web de la CADA : http://www.cada.fr/IMG/pdf/cada_rapport_activite_2015.pdf

[2] Ibid.

[3] Légifrance : https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000031366350

Droits d'auteur réservés. OU (cc) BY-NC-ND

Selon le R2GA[1], et d'après la jurisprudence de la CADA, une demande est qualifiée d'abusive en fonction :

- Du nombre de demandes et du volume de documents demandés
- Du caractère répétitif et systématique des demandes, notamment sur un même sujet
- De la volonté de nuire à l'administration
- Dans le cas d'un contexte tendu entre le demandeur et l'administration saisie

b) La CADA et l'évolution du contexte législatif

Enfin, l'année 2015 est une année où la législation a beaucoup évoluée concernant le droit d'accès aux documents administratifs. La CADA a été sollicité par le gouvernement pour chacun d'eux.

Le premier chantier était celui de la codification de la loi du 17 juillet 1978 dans le livre III du code des relations entre le public et l'administration. Cette codification représente la véritable « *lex generalis* » du droit d'accès aux documents administratifs. Le second chantier a concerné la transposition de la directive 2013/37/UE du Parlement européen et du Conseil du 26 juin 2013, modifiant la directive 2003/98/CE du 17 novembre 2003 concernant la réutilisation des informations du secteur public. Nous en parlerons plus en détail dans le chapitre II, tout comme le dernier chantier qui concerne la loi n° 2016-1321 du 7 octobre 2016 pour une République numérique.

III> LES MODALITÉS D'ACCÈS MATÉRIEL AUX ARCHIVES PUBLIQUES, DANS LE CONTEXTE PAPIER

Au-delà du droit d'accès, les modalités d'accès aux archives publiques sont également encadrées par la loi, toujours dans le Code du patrimoine (A). Cet accès matériel s'effectue le plus généralement dans une salle de lecture prévue à cet effet (B), où le lecteur est guidé dans sa recherche tant par des archivistes que par des outils mis à sa disposition (C). Enfin, pour avoir un regard global sur la communication des archives publiques (papier) en France, nous commenterons certaines données chiffrées en prenant uniquement le cas des Archives départementales (D).

[1] Référentiel général de gestion des archives :

http://references.modernisation.gouv.fr/sites/default/files/Referentiel%20General%20de%20Gestion%20des%20Archives%20R2GA%20-%20octobre%202013.pdf

Droits d'auteur réservés. OU (cc) BY-NC-ND

A> Les dispositions législatives

La communication est la mise à disposition d'archives, dans le sens de document. Ce document peut être papier ou numérique, mais nous nous intéresserons uniquement au support papier dans cette partie.

Si nous revenons à l'article L213-1 du code du patrimoine, celui-ci nous dit que[1] :

> L'accès à ces archives s'exerce dans les conditions définies pour les documents administratifs à l'article L311-9 du code des relations entre le public et l'administration.

L'article cité (L311-9) nous indique que l'accès matériel au document d'archives publiques se fait au choix du demandeur, mais dans les limites techniques de l'administration :

- Par consultation gratuite sur place, sauf si la préservation du document ne le permet pas.
- Par la délivrance d'une copie de même nature aux frais du demandeur, si la reproduction ne nuit pas à la conservation du document.
- Par courrier électronique et sans frais, seulement si le document est disponible sous format électronique.
- Par publication des informations en ligne, sauf si le document n'est communicable qu'à l'intéressé en application de l'article L311-6[2].

Nous nous intéresserons aux deux derniers points dans le chapitre suivant.

B> La salle de lecture

En ce qui concerne la consultation sur place, un document d'archive n'est consultable que dans un environnement contrôlé, donc soit dans une salle de lecture, soit dans un espace de consultation. Nous prenons ici le cas d'une salle de lecture d'un centre d'archives départementales.

[1] Légifrance : https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000006074236

[2] Lié au respect de la vie privée. Ce dernier point a été rajouté suite à loi Lemaire, dont nous parlerons dans le chapitre II.

Droits d'auteur réservés. OU (cc) BY-NC-ND

L'accès aux espaces de consultation est gratuit. Mais cet accès est subordonné à la possession d'une carte de lecteur, établie gratuitement sur présentation d'une pièce d'identité. L'inscription est à renouveler à chaque année civile.

Au lecteur est attribuée une place dans la salle de lecture. Les demandes de communication sont faites par les lecteurs eux-mêmes. Une fois commandés, les documents d'archives sont récupérés dans les magasins pour être apportés en salle de lecture. Généralement, le nombre d'articles communiqué par jour est limité, et il n'est délivré qu'un seul article à la fois. Cela évite de mélanger les documents.

Tout lecteur est soumis à un règlement de la salle de lecture. Celui-ci détaille les conditions générales d'accès à la salle de lecture (notamment le dépôt dans des casiers sécurisés des éléments inutiles pour la consultation des documents). Puis le règlement rappelle les modalités de consultation (délais de communicabilité, commande des documents, conditions de reproduction…). Enfin, sont mentionnées les sanctions encourues par les lecteurs en cas de non-respect du règlement.

C> L'aide à la recherche

Les services proposent également généralement des notices d'aide à la recherche. Ces notices se présentent sous la forme d'un manuel, d'un fascicule ou d'une fiche. Ces aides à la recherche sont fonction de la demande du lectorat. Il s'agit donc souvent de généalogie.

Les inventaires et instruments de recherche sont disponibles en libre accès, soit dans la salle de lecture, soit en salle des inventaires.

Afin de trouver le document que l'on souhaite consulter dans l'immensité des fonds d'archive que conserve un centre d'archives départementales, se référer à un instrument de recherche est indispensable. Il s'agit de la véritable clé d'accès aux fonds d'archives. Ces instruments de recherche sont élaborés par les archivistes suite à un travail d'analyse du fonds.

Il existe depuis 1994 une normalisation de la description archivistique, réalisée par le Conseil international des archives. Il s'agit de la norme ISAD(G) (*International Standard for Archival description - General*)[1]. L'adoption de cette norme facilite la recherche car

[1] CHAVE Isabelle (sous la direction de) et collectif d'auteurs, *Abrégé d'archivistique, principes et pratiques du métier d'archiviste*, Paris, AAF, 2012

Droits d'auteur réservés. OU (cc) BY-NC-ND

elle fournit des informations pertinentes et homogènes d'un fonds à un autre. Les instruments de recherche ISAD(G) comportent un grand nombre de niveaux et sont structurés de telle manière qu'il existe un lien hiérarchique entre les descriptions, du général au particulier.

La norme décline une liste de 26 éléments de description au travers de 7 zones. Parmi tous ces éléments, 6 sont indispensables : il s'agit de la référence (identifiant), de l'intitulé, des dates extrêmes, du niveau de description (fonds, sous-fonds…), de l'importance matérielle (ml) et enfin du nom du producteur[1].

Un instrument de recherche papier classique comporte lui aussi 7 zones, qui correspondent aux mêmes zones décrites dans la norme ISAD(G) : l'introduction, la bibliographie et les sources complémentaires, le sommaire, le corps de l'instrument de recherche, les annexes, l'index alphabétique et enfin la table des matières[2].

Le personnel de l'accueil et des salles de consultation peut guider les lecteurs dans la consultation des inventaires, fichiers et usuels disponibles, mais il n'a pas vocation à se substituer à eux dans leurs recherches. De même, les services d'archive ne sont pas tenus de faire des recherches par correspondance. Des recherches de ce type peuvent être menées en fonction de la disponibilité des agents. Plus généralement, l'archiviste fournira au lecteur une liste de cotes d'archives où celui-ci sera susceptible de trouver les informations qu'il cherche. Il devra donc ensuite se déplacer en salle de lecture pour les consulter. Certains services peuvent fournir un service de recherche payant, étant donné qu'il s'agit d'une mission qui dépasse les obligations de service public.

D> Les chiffres de 2015

Chaque année, les services d'archives (nationales, régionales, départementales et communales) doivent fournir au SIAF (Service Interministériel des Archives de France) une grille de chiffres, concernant 11 domaines :

- Le personnel et les budgets
- L'accroissement des fonds
- Les instruments de recherche
- La conservation et restauration

[1] CHAVE Isabelle (sous la direction de) et collectif d'auteurs, *Abrégé d'archivistique, principes et pratiques du métier d'archiviste*, Paris, AAF, 2012

[2] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

- Le contrôle scientifique et technique (CST)
- L'occupation de l'espace
- La numérisation
- La mise en ligne
- La communication
- La consultation en ligne
- Les expositions et animations

Celle qui nous intéresse, la rubrique communication (papier) est divisée en plusieurs catégories : le nombre de lecteurs inscrits (et proportions de lecteurs généalogistes, chercheurs…), le nombre de séances de travail, le nombre de communications, le nombre de recherches par correspondance, le nombre de dérogations instruites (ainsi que le nombre d'articles accordés et refusés). Nous pouvons relever quelques chiffres intéressants concernant l'année 2015[1] (tableau le plus récent).

Au vu de ce tableau, le premier constat que nous pouvons faire est que les chiffres sont très hétérogènes en fonction des AD. Le nombre de lecteurs inscrits varie entre 6 993 pour Paris et 126 pour Mayotte. Les publics majoritaires sont sans surprise les généalogistes (ils sont en moyenne 399, pour 222 chercheurs et 160 lecteurs venus pour des recherches individuelles ou administratives).

Au-delà du lectorat, ce qui nous intéresse le plus vis-à-vis de ce mémoire sont les chiffres de séances de travail, de communications et de recherches par correspondance.

Au total, 1 356 875 documents ont été communiqués en 2015 dans l'ensemble des AD, pour 305 412 séances de travail. Nous pouvons remarquer que le chiffre de recherches par correspondance est élevé, 130 545 pour l'ensemble des AD. Au regard du nombre de communication, cela correspond tout de même à 10%. Encore une fois, les chiffres sont vraiment disparates entre les départements. On pourrait penser que cela est lié à la démographie du territoire. Exemple : 55 297 communications à Paris (2 240 000 habitants), pour 2 718 en Corse du Sud (150 000 habitants). Mais ce n'est pas forcément lié. Exemple : les Yvelines font 7 716 communications (1 400 000 habitants) alors que l'Allier en fait 35 060 (343 000 habitants).

[1] Portail France Archives : https://francearchives.fr/article/37978

Droits d'auteur réservés. OU (cc) BY-NC-ND

Ces chiffres montrent qu'il existe un véritable ancrage territorial. Un centre d'AD peut être plus intéressant et plus visité qu'un autre car il conserve par exemple une plus riche collection, chaque département adoptant sa propre stratégie pour attirer le lecteur et faire vivre ses archives.

Dans ce premier chapitre nous avons étudié les tenants et aboutissants de l'accès aux archives publiques, dans un contexte papier. Nous cherchons désormais à savoir quel impact le numérique a-t-il eu sur cet accès ?

Droits d'auteur réservés. OU (cc) BY-NC-ND

CHAPITRE II : LA RÉVOLUTION NUMÉRIQUE OU LE PASSAGE DE LA COMMUNICATION À LA DIFFUSION

Dans ce chapitre nous allons voir que le numérique a rebattu les cartes de l'accès aux archives publiques. Aujourd'hui, une démarche de diffusion tend à primer sur celle de la communication. D'une manière proactive, les services d'archives publics versent une grande quantité d'archives publiques en ligne et deviennent ainsi librement accessibles.

Nous allons donc tenter de répondre à trois questions : comment, quoi et qui ? Comment s'est produit ce renversement, à quelles archives avons-nous accès en ligne et enfin qui sont les acteurs ?

La règlementation a dû s'adapter à ce nouveau contexte d'ouverture des données et aux nouvelles pratiques qu'offre l'accès en ligne (I). De nouvelles formes d'accès ont vu le jour à travers les portails open data et les sites web des services publics d'archives, mettant à disposition de l'internaute une masse importante de documents (II). Enfin, nous constaterons que les services publics d'archives ont perdu leur monopole de communication des archives publiques, car de nouveaux acteurs sont apparus (III).

I> UN NOUVEAU CONTEXTE – TANT LÉGISLATIF QUE SOCIÉTAL – FAVORABLE À L'OUVERTURE DES DONNÉES PUBLIQUES

Le numérique ayant profondément bouleversé le rapport à l'accès aux archives publiques, cette première partie vise à répondre à la question du « comment ». Comment ce bouleversement a-t-il eu lieu ? Pourquoi ouvrir les données ? Comment l'open data s'est-il développé en France (A) ?

Enfin nous allons voir qu'une législation très récente est venue encadrer ce nouveau contexte d'ouverture des données (B).

Droits d'auteur réservés. OU (cc) BY-NC-ND

A>L'ouverture des données

1) Le mouvement open data

a) Principes

Le mouvement open data a pour objectif la mise à disposition de données en provenance des institutions (publiques et privées) et à destination des citoyens[1]. L'administration collecte et produit énormément de données dans le cadre de son activité, et le but de l'open data est que l'administration publique ouvre ces données à la collectivité.

Étymologiquement, data vient du latin et signifie « choses données[2] ». Ce mot, conservé par les anglosaxons, fait aujourd'hui référence à des informations numériques. En français nous parlons simplement de « données », mais nous avons conservé des expressions anglophones comme par exemple *open data*, *data mining*, *big data*, *data science* etc. Dans notre monde actuel, les données sont omniprésentes et représentent des informations considérables par leur volume et par leurs valeurs : la météo, les cartes routières, les horaires de transport en commun, les résultats sportifs, les résultats des élections, le PIB etc. Ainsi, les données sont indispensables. Dans notre monde contemporain, la donnée est de l'or noir. Obtenir, traiter et diffuser de la donnée est devenu un enjeu stratégique car celle-ci est source de profits.

Avec l'open data, les données sont ainsi ouvertes et libres. Elles sont encadrées par un choix de licences qui décrivent ce qu'il est possible de faire avec ces données : utilisation, réutilisation, conditions particulières...

b) Objectifs

Nous pouvons détailler les principaux objectifs de l'open data[3] :

- La transparence

Il s'agit peut-être de l'objectif principal de l'open data. En exposant leurs données dans les formats plus facilement consommables que des registres, les institutions accomplissent de manière plus formelle leur devoir d'ouverture. En effet, lorsque la

[1] BOURCIER Danièle, FILIPPI Primavera de, *Open data & Big data : nouveaux défis pour la vie privée*, Paris, Mare & Martin

[2] Wiktionnaire : https://fr.wiktionary.org/wiki/data

[3] GOUIGOUX Jean-Philippe, *Open data : consommation, traitement, analyse et visualisation de la donnée publique*, St Herblain, Editions ENI, 2016

Droits d'auteur réservés. OU (cc) BY-NC-ND

donnée publique existe sur format papier et se trouve localisée dans des centres précis, sa consommation n'est pas évidente ni facile. L'approche open data permet de renouveler et approfondir cette ouverture des institutions aux citoyens, en leur offrant une façon plus simple et plus efficace de consommer de la donnée. De plus, la transparence de l'action publique est un enjeu de société actuel, comme nous avons pu le constater suite aux dernières élections présidentielles.

- La valorisation économique

Au-delà de cet objectif de transparence, qui est plutôt désintéressé, l'open data permet également de créer de la richesse. L'objectif économique est donc important. La mise à disposition de données ouvertes permet d'inventer des usages de valorisation et des nouveaux logiciels qui faciliteront la vie des utilisateurs. Nous en parlerons plus en détails dans le chapitre suivant. L'open data a donc permis à des entreprises de se créer.

De plus, comme l'open data est souvent de portée locale, la réutilisation des données a l'avantage de favoriser les PME et start-ups locales, qui créeront des usages pour les locaux également. Ceci permet de localiser la production de valeur.

- Le service rendu aux citoyens

A leur tour, ces utilisations valorisant la donnée publique, qu'elles soient gratuites ou commerciales, participent à un troisième objectif, à savoir un meilleur service public rendu au citoyen, que cela soit directement par la consommation des données ouvertes à la source (sur un portail fourni par la collectivité par exemple) ou indirectement par les applications construites sur cette donnée.

Le service supplémentaire pour les citoyens est un aspect positif recherché par l'open data. La mise à disposition de données doit toutefois être bien pensée pour favoriser la réutilisation, mais aussi pour être pertinente par rapport aux besoins des citoyens.

- La valorisation interne

L'utilisateur n'est pas forcément le citoyen, en effet l'administration elle-même peut jouir de l'open data. Les collectivités consomment également leurs propres données, qui ont été qualifiées et nettoyées. De plus, le fait de proposer de la donnée ouverte devient de plus en plus un facteur d'attractivité des territoires et des villes. La valorisation open data est donc non seulement celle de la donnée mais aussi celle du fournisseur de cette donnée.

Droits d'auteur réservés. OU (cc) BY-NC-ND

Enfin, notons tout de même que d'autres objectifs – moins nobles – peuvent entrer en jeu. Il s'agit par exemple de l'affichage politique. L'open data peut être un moyen de se démarquer par rapport à un autre candidat. Mais si c'est bien cette orientation qui est prise, elle est rapidement visible. En effet, lorsqu'aucun jeu de données n'est pas rafraichi et que toutes les données datent de la mise en ligne initiale, il n'y a que peu de doute sur la volonté d'affichage politique. A l'inverse, un site en enrichissement permanent avec des données à jour est la marque d'un réel engagement dans la démarche.

2) Historique du mouvement

a) Les anglo-saxons à la barre

L'open data n'est ni une technologie ni un simple outil, mais bel et bien un mouvement politique, il est essentiel d'expliquer sa naissance et son historique.

Les premières apparitions du concept d'open data datent du premier mandat de Barack Obama aux EU[1]. Le principe est de rendre disponibles les données utilisées par l'administration publique dans le cadre de son activité pour une réutilisation par d'autres entités que l'Etat, typiquement les entreprises et les citoyens. Il y a là une bivalence de l'enjeu politique : servir à la fois à améliorer la compétitivité des entreprises et fournir un niveau de transparence de l'action publique plus élevé au citoyen.

Mais c'est bien dans un premier temps la transparence qui a été évoquée comme objectif du mouvement open data. Dès 2007, la loi dite Open Government Act posait les principes d'un accès libre à l'information publique[2]. C'est lors de sa campagne électorale que Barack Obama a annoncé son intention de « renforcer la démocratie et son efficacité[3] » par la mise à disposition de plus de données collectées par l'État. L'Open Government Act Initiative est lancée véritablement en janvier 2009.

Puis c'est Vivek Kundra, le premier DSI de l'État (*Chief Information Officer*) qui constatera que l'ouverture des données a donné lieu à de nombreux projets d'utilisation en lien avec des enjeux économiques[4]. Des start-ups se sont en effet créées pour valoriser cette donnée jusque-là inaccessible et fournir des services innovants.

---

[1] GOUIGOUX Jean-Philippe, *Open data : consommation, traitement, analyse et visualisation de la donnée publique*, St Herblain, Editions ENI, 2016

[2] Ibid.

[3] Ibid.

[4] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

La manifestation la plus concrète de la politique open data est le lancement du site data.gov, dès 2009, qui regroupe en un seul point d'accès tous les jeux de données fournis par le gouvernement des EU[1].

Le Royaume-Uni lance dès 2010 un projet calqué sur celui des EU, avec une approche très structurée à la manière de Tim Bernes-Lee[2]. L'équivalent britannique de data.gov, à savoir data.gov.uk, présente tous les jeux de données dans un catalogue informatique lui-même accessible comme un jeu de données. Ceci facilite clairement l'exploration des données disponibles. L'importance de la sémantique était déjà constatée.

De plus, le portail accompagne les données de nombreuses métadonnées. Elles permettant à l'utilisateur de mieux appréhender et de mieux consommer la donnée brute. Par exemple, le portail va fournir dans un jeu de données des nombres, accompagnés de l'unité dans laquelle ce nombre est exprimé, sa période de valeur dans le temps, ses relations aux autres jeux de données, etc.

b) Une initiative poursuivie en France

Les pionniers de l'open data en France sont les villes de Rennes et Paris.

Dès 2010, la ville de Rennes a ouvert de la donnée au public[3]. Ces données concernaient les transports (comme c'est souvent le cas). Le portail Data Rennes Métropole[4] est lancé en juillet 2010. L'approche de Rennes était dès le début extrêmement collaborative, avec l'inclusion des associations comme la Cantine Numérique[5], et une volonté forte de prendre en compte les retours des utilisateurs de la donnée pour améliorer en continu sa mise à disposition.

Quelques mois plus tard, c'est la ville de Paris qui ouvre son portail en janvier 2011, et qui ne cessera de s'enrichir de nouvelles données et fonctionnalités[6]. La Ville de Paris est encore aujourd'hui un des membres les plus actifs de l'open data, avec de nombreuses

[1] Portail open data des EU : https://www.data.gov/

[2] Jean-Philippe Gouigoux, op.cit.

[3] Ibid.

[4] Portail open data de la ville de Rennes : https://data.rennesmetropole.fr/page/home/

[5] Structure dédiée au numérique sur Rennes Métropole. C'est un lieu d'échanges et de débats, qui organise régulièrement des événements autour du numérique.

[6] Jean-Philippe Gouigoux, op.cit.

Droits d'auteur réservés. OU (cc) BY-NC-ND

améliorations techniques du portail, ce qui permet une consommation plus aisée et une plus forte valorisation.

Le cas de Bordeaux est également intéressant. L'équipe en charge du projet s'est rendu compte qu'en réalité le premier utilisateur des données était son producteur. En effet, le fait d'ouvrir des données oblige à les nettoyer, rendre leur accès plus simple et, au final, amène les mêmes personnes qui les ont ouvertes à se rendre compte de nouveaux usages possibles. Un exemple parmi tant d'autres de ce phénomène est que la première réutilisation des données sur les essences des arbres (initialement à destination des personnes allergiques aux pollens) a été réalisée en interne par les services d'entretien des jardins, qui ont pu optimiser leur tournée de traitement des espèces végétales grâce à ce référentiel[1].

En France, l'État est arrivé sur l'open data un peu après les villes pionnières, et a entrepris de se positionner en entrepôt central de la donnée. Un décret de février 2011 crée ETALAB, un service dédié à la donnée publique et rattaché au Premier ministre. Le Premier ministre François Fillon indique que « cette initiative contribuera donc à renforcer la transparence de l'action de l'État. Elle ouvrira le champ à une discussion avec les citoyens à qui les institutions – celles de l'État, mais aussi celles des collectivités locales – doivent rendre des comptes[2]». Nicolas Sarkozy revient lui aussi sur le lancement du portail français : « Il ne sert à rien de résister à cette transparence. Il vaut beaucoup mieux jouer le jeu de l'accès aux données publiques que de résister. Il n'y a pas de secrets d'État. Ce n'est pas possible ! ».[3]

Le portail data.gouv.fr ouvre en décembre de la même année, avec plusieurs centaines de milliers de jeux de données.

En 2011 ont ouvert les portails de Toulouse, Nantes, Montpellier, ainsi que la Gironde et la Saône-et-Loire, parmi les plus actifs à l'époque. Ils ont été rejoints en 2012 puis en 2013 par une quantité d'autres villes, départements et régions. Aujourd'hui, toutes les grandes villes de France ont leur portail de données. Avec la consolidation des villes en groupements de communes, les agglomérations sont en train de devenir le porteur idéal

[1] GOUIGOUX Jean-Philippe, *Open data : consommation, traitement, analyse et visualisation de la donnée publique*, St Herblain, Editions ENI, 2016

[2] CHIGNARD Simon, *Open data : comprendre l'ouverture des données publiques*, Limoges, Editions Fyp, 2012

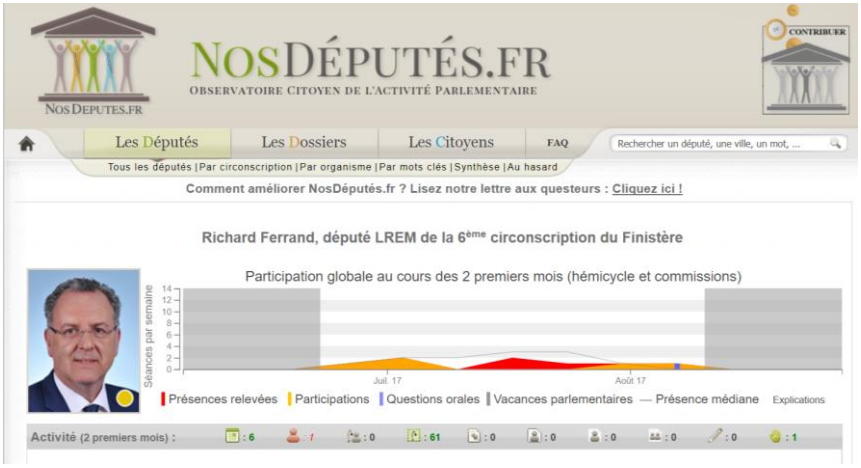
[3] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

des données publiques, car leur regroupement permet des économies, tout en restant à une taille idéale pour une approche locale de la donnée.

c) Un mouvement soutenu et encouragé par des associations

Enfin, de nombreuses associations se sont emparées du sujet et ont favorisé la valorisation de l'open data en France. Les deux plus connues sont LiberTIC[1] et Regards Citoyens[2]. Les membres de cette dernière association souhaitent « proposer un accès simplifié au fonctionnement des institutions démocratiques ». Ils créent notamment le site nosdéputés.fr. Ce site nous donne quantité d'informations sur les députés : présences relevées à l'Assemblée nationale, nombre de participations, nombre de questions orales etc. et permet au citoyen de se faire une idée de la représentation que le député lui apporte. Ci-dessous une capture d'écran du site[3], nous donnant des informations sur Richard Ferrand, député de la 6ème circonscription du Finistère.



Enfin, les collectivités productrices d'open data ont elles aussi constitué une association nommée « open data France », et qui a pour but de soutenir le mouvement open data en regroupant les participants et leur fournissant du soutien. Celui-ci se produit sous forme de mutualisation de bonnes pratiques, mais aussi de par une représentation auprès d'autres instances. L'association open data France est particulièrement active est très bien implantée parmi les personnes en charge de l'open data dans les collectivités motrices sur le sujet[4].

[1] Site Web de l'association LiberTIC : https://libertic.wordpress.com/libertic/

[2] Site Web de l'association Regards citoyens : http://www.regardscitoyens.org/

[3] Site Web nosdeputes.fr : https://www.nosdeputes.fr/richard-ferrand, consulté le 15/08/17

[4] Site Web open data France : http://www.opendatafrance.net/

Droits d'auteur réservés. OU (cc) BY-NC-ND

B> Évolutions législatives et règlementaires

La France dispose d'un terreau favorable à l'apparition du mouvement open data, dans le sens où la loi sur le droit à l'information impose dès 1978 aux collectivités locales ainsi qu'à l'Etat de mettre à disposition des citoyens en faisant la demande les données qui les concernent. Il n'y avait donc qu'un pas à réaliser pour mettre à disposition ces données sans attendre une requête de la part du citoyen, et sous une forme plus facilement utilisable.

Depuis 2015, du plan d'action national pour une action publique et transparente et collaborative (1), en passant par la loi Valter (2) et l'incontournable loi Lemaire « pour une République numérique » (3), de nombreux décrets, arrêtés, lois et directives européennes comme nationales ont fait progressivement avancer la cause de l'open data. Malgré les réticences de certains groupes craignant une trop forte transparence, la France se situe honorablement dans le classement international sur l'ouverture des données publiques. En effet la France se positionne à la troisième place du classement Global open data Index[1] ce qui traduit une forte volonté de sa part dans la poursuite de ce mouvement.

1) Le plan d'action national pour une action publique transparente et collaborative

La France est présidente du Partenariat pour un gouvernement ouvert[2] (PGO) depuis l'automne 2016 pour une durée d'un an. Dans le cadre de ce partenariat, François Hollande a souhaité mettre en place un plan d'action national pour une action publique transparente et collaborative. En effet, le Président de la République française est conscient que « trop de concitoyens en viennent à douter des institutions, de leurs dirigeants, de la capacité même des pouvoirs publics à agir face aux grands défis de notre temps[3] ».

C'est pourquoi, selon lui, un des principes fondamentaux à instaurer est l'ouverture de l'action publique, « afin qu'elle ne dépende pas seulement de l'État et des collectivités

[1] http://index.okfn.org/place/

[2] Ce partenariat regroupe aujourd'hui 75 pays et des centaines d'organisations de la société civile qui agissent à travers le monde pour la transparence de l'action publique, pour sa co-construction avec la société civile et les citoyens et pour l'innovation démocratique : https://www.etalab.gouv.fr/ogp

[3] http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/pgo_plan_action_france_2015-2017_fr.pdf

Droits d'auteur réservés. OU (cc) BY-NC-ND

publiques, mais que la société civile elle-même puisse y avoir sa part et contribuer à la construction du bien commun[1] ».

L'ouverture des données publiques constitue en effet un levier pertinent pour renforcer la qualité du lien entre les responsables publics et les citoyens. Comme le souligne le Conseil d'État, elle « ouvre à tout citoyen ou à tout groupement un droit de regard sur les moyens et les résultats des politiques publiques, lui permettant de dénoncer des dysfonctionnements, voire de contribuer à leur résolution[2] ».

En cohérence avec sa politique d'ouverture des données publiques, le gouvernement entend prolonger son effort[3] :

- En travaillant au développement et à la mise à disposition des ressources technologiques ouvertes, parmi lesquelles les données de référence, mais aussi les modèles de la décision et de l'action publique.
- En développant une stratégie technologique d'État selon la logique du « gouvernement comme plateforme ». Celle-ci permet d'ouvrir les ressources informatiques, facilitant leur appropriation par tous et renforçant leur impact.
- En poursuivant le développement de l'administration numérique et de la dématérialisation des services publics.

L'engagement n°15 du plan d'action[4] nous semble le plus intéressant vis-à-vis de notre sujet :

> Engagement 15 : renforcer la politique d'ouverture et de circulation des données
>
> Cette politique essentielle est considérée à la fois comme un ressort de vitalité démocratique, une stratégie d'aide à l'innovation économique et sociale, et un levier pour la modernisation de l'action publique.

[1] http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/pgo_plan_action_france_2015-2017_fr.pdf

[2] Conseil d'État, Le numérique et les droits fondamentaux, Étude annuelle 2014.

[3] http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/pgo_plan_action_france_2015-2017_fr.pdf

[4] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

Nous retrouvons bien dans cet engagement tous les objectifs de l'open data précédemment décrits.

2) La directive européenne PSI et la loi Valter

La directive européenne PSI (*public sector information*) de 2003 porte sur la réutilisation des informations détenues par le secteur public. Elle définit également les domaines où la réutilisation est autorisée. La révision en 2013 de la directive PSI exige que tout document entrant dans son champ d'application soit rendu accessible pour pouvoir être réutilisé, tant pour des usages commerciaux que non commerciaux. C'est l'affirmation d'un véritable droit à réutilisation des données du secteur public[1].

La France transpose cette directive dans le droit français avec la loi n° 2015-1779 du 28 décembre 2015 relative à la gratuité et aux modalités de la réutilisation des informations du secteur public, dite loi Valter[2].

Depuis cette loi, la réutilisation est donc un droit, et concerne aussi les services d'archives. Bruno Ricard, sous-directeur au SIAF, note deux exceptions à ce principe[3] : tout d'abord, la loi autorise les services culturels à demander des redevances si les fichiers sont non nativement numériques et sont issus d'une numérisation. Le tarif de numérisation est lui-même encadré par la loi. D'autre part, les services culturels ont la possibilité de conclure des partenariats avec des entreprises privées chargées d'effectuer la numérisation. Celles-ci obtiendraient alors des droits d'exploitation pendant une certaine durée (quinze ans au maximum).

La loi Valter demande également aux administrations qui diffusent leurs données de le faire « si possible » dans « un standard ouvert et aisément réutilisable, c'est-à-dire lisible par une machine[4] ». Au départ, la mention du « si possible » n'existait pas[5], mais le gouvernement l'a rajoutée pour encourager les administrations à poursuivre ce mouvement, sans leur imposer trop de complications.

---

[1] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:175:FULL:FR:PDF

[2] Clotilde Valter est alors Secrétaire d'Etat chargée de la réforme de l'Etat et de la simplification.

[3] Forum des archivistes 2016 à Troyes : session archives accessibles/archives ouvertes ? « open data et réutilisation : un droit enfin stabilisé ? » : https://chartes.hypotheses.org/853

[4] Légifrance : https://www.legifrance.gouv.fr/eli/loi/2015/12/28/PRMX1515110L/jo

[5] Xavier Berne, « La loi sur la gratuité des données publiques définitivement adoptée par le Parlement » : https://www.nextinpact.com/news/97778-la-loi-sur-gratuite-donnees-publiques-definitivement-adoptee-par-parlement.htm

Droits d'auteur réservés. OU (cc) BY-NC-ND

Une nouvelle fois, nous assistons à un renversement de paradigme. Auparavant la réutilisation pouvait donner lieu au versement de redevances, aujourd'hui cette réutilisation est gratuite par défaut.

Notons que seuls certains services dérogent à ce principe, comme par exemple l'IGN[1], Météo France ou encore le SHOM[2], car ces services qui ont des missions très particulière « doivent couvrir par des recettes propres une part substantielle des couts liés à l'accomplissement de ses missions de service public[3] ». Dans ce cas, le coût de la redevance est fixé et ne doit pas dépasser les coûts de production et de diffusion.

3) La loi Lemaire, dite « pour une République numérique »

La loi pour une République numérique a été promulguée le 7 octobre 2016. La loi entend encourager l'innovation et l'économie numérique, promouvoir une société numérique protectrice et garantir l'accès de tous au numérique[4].

Le 10 octobre 2016, Axelle Lemaire déclare : « Nous avons, ensemble, voulu une loi numérique porteuse d'un souffle de transparence, d'ouverture et d'innovation. Nous avons bâti un socle pour garantir les droits de chacun et la confiance de tous dans l'économie numérique. Nous avons exprimé l'exigence commune d'accompagner tous nos concitoyens, dans tous les territoires, à la transition numérique en cours, afin que personne ne soit laissé à l'arrière des avancées économiques et sociales dont elle porte la promesse[5] ». L'objectif de cette loi serait de « faire de la France une "République numérique" dont la devise serait : Liberté d'innover, Égalité des droits, Fraternité d'un numérique accessible à

[1] Institut national de l'information géographique et forestière

[2] Service hydrographique et océanographique de la Marine

[3] Légifrance : https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000031701525&categorieLien=id

[4] Secrétariat d'état au numérique, « La loi pour une République numérique, préparer notre pays aux enjeux de la transition numérique » : https://www.economie.gouv.fr/files/files/PDF/DP-RN-160111.PDF

[5] https://www.economie.gouv.fr/files/files/PDF/DP_LoiNumerique.pdf

Droits d'auteur réservés. OU (cc) BY-NC-ND

tous et Exemplarité d'un État qui se modernise[1] ».

La Secrétaire d'État chargée du numérique et de l'innovation souhaite ainsi « libérer l'innovation en faisant circuler les informations et les savoirs, pour armer la France face aux enjeux globaux de l'économie de la donnée[2] ». Selon elle, « l'open data par défaut des administrations publiques ouvrira une large partie des données publiques, et l'ouverture de données d'intérêt général d'acteurs privés chargés d'une mission de service public, bénéficiera à l'ensemble de la société [3]».

C'est la première partie de cette loi qui nous intéresse vis-à-vis de notre sujet. Celui-ci cherche à favoriser la « circulation des données et du savoir » à travers l'ouverture des données publiques par défaut (a), la création d'un service public de la donnée (b), l'ouverture des données d'intérêt général (c) et le libre accès aux écrits de la recherche publique (d).

a) L'ouverture des données par défaut

La loi consacre l'ouverture des données par défaut. L'open data devient la règle et non plus l'exception. Les administrations doivent désormais publier en ligne et dans un standard ouvert leurs principaux documents, leurs codes sources, leurs bases de données ainsi que leurs données qui présentent un intérêt économique, social, sanitaire ou environnemental[4]. Sont concernés les administrations d'État, les collectivités locales de plus de 3 500 habitants, les établissements publics et les organismes privés chargés d'un service public.

Pour illustrer ce point, voici quelques extraits de cette loi, qui a été intégrée au CRPA[5] :

Article L 312-1-1 :

---

[1] https://www.senat.fr/espace_presse/actualites/201603/projet_de_loi_pour_une_republique_numerique.html

[2] 28 septembre 2016, communiqué de presse d'Axelle Lemaire intitulé « la loi pour une République numérique adoptée à l'unanimité au Sénat : http://proxy-pubminefi.diffusion.finances.gouv.fr/pub/document/18/21499.pdf

[3] Ibid.

[4] http://www.vie-publique.fr/actualite/dossier/loi-internet/republique-numerique-que-change-loi-du-7-octobre-2016.html

[5] Légifrance : https://www.legifrance.gouv.fr/affichCode.do;jsessionid=CAB94BE3DF8ADA1519EAA9099EA78111.tpdila07v_3?idSectionTA=LEGISCTA000031367737&cidTexte=LEGITEXT000031366350&dateTexte=20170831

Droits d'auteur réservés. OU (cc) BY-NC-ND

> « Les administrations […] publient en ligne les documents administratifs suivants :
>
> 1) Les documents qu'elles communiquent en application des procédures prévues au présent titre, ainsi que leurs versions mises à jour ;
> 2) Les documents qui figurent dans le répertoire mentionné au premier alinéa de l'article L. 322-6[1] ;
> 3) Les bases de données, mises à jour de façon régulière, qu'elles produisent ou qu'elles reçoivent et qui ne font pas l'objet d'une diffusion publique par ailleurs ;
> 4) Les données, mises à jour de façon régulière, dont la publication présente un intérêt économique, social, sanitaire ou environnemental.
>
> Le présent article ne s'applique pas aux collectivités territoriales de moins de 3 500 habitants. »

Notons que les algorithmes publics et leurs règles sont également concernés.

L'article L300-4 fait mention des standards de publication :

> « Toute mise à disposition effectuée sous forme électronique en application du présent livre se fait dans un standard ouvert, aisément réutilisable et exploitable par un système de traitement automatisé ».

Par cet article, l'état cherche à encourager la réutilisation des données publiques. Les fichiers PDF sont donc bannis car inexploitables. Les déclarations d'intérêts des parlementaires devront donc être accessibles sous un autre format qui permettra leur réutilisation.

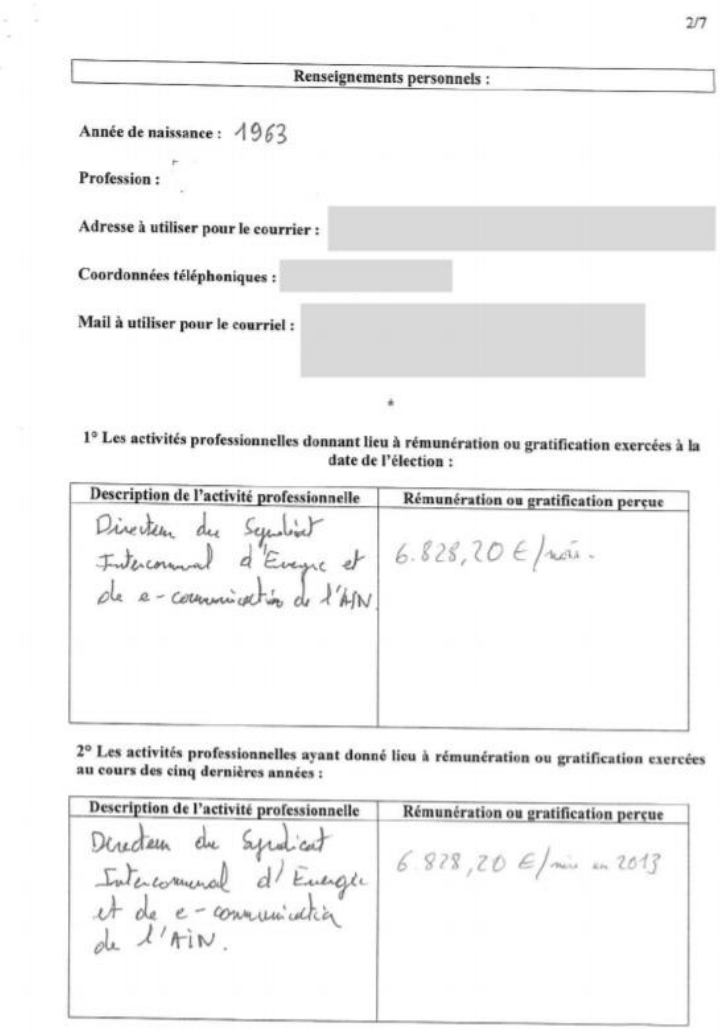
Ci-dessous un exemple PDF de la déclaration d'intérêts du sénateur de l'Ain Patrick Chaize[2]. Nous constatons qu'une réutilisation de ces informations serait laborieuse.

[1] L 322-6 : Les administrations qui produisent ou détiennent des informations publiques tiennent à la disposition des usagers un répertoire des principaux documents dans lesquels ces informations figurent. Elles publient chaque année une version mise à jour de ce répertoire.

[2] Toutes les déclarations sont disponibles sur le site de la Haute autorité pour la transparence de la vie publique : http://www.hatvp.fr/livraison/dossiers/chaize-patrick-dia-senateur-01.pdf

Droits d'auteur réservés. OU (cc) BY-NC-ND

Le réutilisateur doit créer par exemple un fichier excel pour y retranscrire toutes les données. De plus, utiliser un autre format permettrait d'éviter d'avoir à déchiffrer l'écriture des parlementaires.



2/7

Renseignements personnels :

Année de naissance : 1963

Profession :

Adresse à utiliser pour le courrier :

Coordonnées téléphoniques :

Mail à utiliser pour le courriel :

1° Les activités professionnelles donnant lieu à rémunération ou gratification exercées à la date de l'élection :

Description de l'activité professionnelle	Rémunération ou gratification perçue
Directeur du Syndicat Intercommunal d'Energie et de e-communication de l'Ain	6.828,20 €/mois.

2° Les activités professionnelles ayant donné lieu à rémunération ou gratification exercées au cours des cinq dernières années :

Description de l'activité professionnelle	Rémunération ou gratification perçue
Directeur du Syndicat Intercommunal d'Energie et de e-communication de l'Ain.	6.828,20 €/mois en 2013

L'article L 311-1 quant à lui est modifié, auparavant, il indiquait ceci[1] :

> « Sous réserve des dispositions des articles L. 311-5 et L. 311-6, les administrations mentionnées à l'article L. 300-2 sont tenues de communiquer les documents administratifs qu'elles détiennent aux personnes qui en font la demande, dans les conditions prévues par le présent livre ».

[1] Légifrance : https://www.legifrance.gouv.fr/affichCode.do;jsessionid=812EF80F0B81EB5F788AB7A633ED1756.tpdila07v_3?idSectionTA=LEGISCTA000031367696&cidTexte=LEGITEXT000031366350&dateTexte=20170831

Droits d'auteur réservés. OU (cc) BY-NC-ND

Suite à la loi Lemaire, après le mot : « tenues », sont insérés les mots : « de publier en ligne ou ».

La loi indique également que la diffusion de certains documents ne pourra se faire que « sous réserve d'anonymisation ou d'occultation » des mentions touchant notamment à la vie privée. Cette diffusion est exclue si elle engage la « sécurité des systèmes d'information des administrations[1] ».

Les pouvoirs de la CADA et de la CNIL sont également renforcés par cette loi. Auparavant, la CADA pouvait être saisie pour refus de « communication » de documents administratifs. Maintenant, elle pourra également être saisie pour refus de de « publication » d'informations publiques. Tout citoyen peut donc saisir la CADA si une administration ne met pas en ligne ses principales données.

Bien sûr, la loi met en place un calendrier. On ne peut pas imposer aux administrations la mise en ligne de leurs données en un temps réduit. Ainsi d'ici 2018, les principaux documents administratifs devront être publiés en ligne. Les administrations seront accompagnées dans ce processus notamment par l'association Open data France.

b) La création d'un service public de la donnée

La loi prévoit également la mise en place d'un service public de la donnée. Ce service est chargé de faciliter la réutilisation des bases de données de l'État. Il regroupera des grandes bases de données de référence comme par exemple le répertoire SIRENE des entreprises, le cadastre ou encore d'autres données INSEE[2].

c) L'ouverture des données d'intérêt général

Une nouvelle notion apparait avec la loi, celle des données d'intérêt général. Qu'est ce qui peut être qualifié d'intérêt général ? Ce sont par exemple les données des délégations de service public (transport, eau, énergies, déchets, etc.), les données relatives aux subventions publiques ou encore les données de consommation d'énergie[3].

---

[1] Légifrance : https://www.legifrance.gouv.fr/affichCode.do;jsessionid=812EF80F0B81EB5F788AB7A633ED1756.tpdila07v_3?idSectionTA=LEGISCTA000031367696&cidTexte=LEGITEXT000031366350&dateTexte=20170831

[2] Secrétariat d'état au numérique, « La loi pour une République numérique, préparer notre pays aux enjeux de la transition numérique » : https://www.economie.gouv.fr/files/files/PDF/DP-RN-160111.PDF

[3] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

d) Le libre accès aux travaux de la recherche publique et l'exploration des données

Lorsque des travaux de recherche sont financés à plus de moitié par des fonds publics, les chercheurs pourront les déposer en libre accès en ligne. C'est une mesure qui permet de faciliter les résultats de recherche afin de concurrencer les éditeurs de publications scientifiques[1].

Nous venons de voir dans cette partie qu'une réglementation très récente est venue encadrer l'ouverture des données publiques et leur diffusion en ligne. Changeons maintenant de point de vue et prenons la place du citoyen : comment avoir accès à ces données ?

II> LES NOUVELLES FORMES D'ACCÈS

Le numérique engendre de nouvelles formes d'accès aux archives et données publiques. De quelle manière le citoyen accède-t-il aux données publiques en ligne ? Concrètement, à quelles archives ou données publiques a-t-il accès ?

Cet accès se matérialise sous la forme de portails en ce qui concerne l'open data (A) et sous la forme de sites internet en ce qui concerne les services publics d'archives tels que les AM, AD ou AN. Nous prendrons ici le cas des AD qui nous semble le plus démonstratif (B).

A> Les portails open data

Comme nous l'avons signalé, l'open data est une démarche récente qui s'inscrit dans l'ère du temps (transparence, stimulation de l'économie, création de nouveaux services en lien avec les nouvelles technologies, écologie, etc.). Nous verrons que cette démarche, qui implique différents acteurs, est règlementée par le biais de licences (1). Nombreux sont ceux à jouer le jeu, tant l'État que des villes, ce qui favorise l'émergence de nombreux portails aux caractéristiques très élaborées (2). Ceux-ci donnent accès à quantité de jeux de données et d'informations publiques, répondant aux objectifs précédemment cités (3).

[1] Secrétariat d'état au numérique, « La loi pour une République numérique, préparer notre pays aux enjeux de la transition numérique » : https://www.economie.gouv.fr/files/files/PDF/DP-RN-160111.PDF

Droits d'auteur réservés. OU (cc) BY-NC-ND

1) Enjeux, acteurs et licences

a) Les enjeux, pourquoi ouvrir les données ?

Au-delà de l'obligation légale, reconnue par le droit français, de mettre à disposition les données publiques, la volonté pour une communauté de mettre en œuvre une démarche d'open data et d'ouvrir un portail peut être guidée par des objectifs variés. En effet, comme nous l'avons déjà signalé en début de ce chapitre, les champs de l'open data sont larges, de l'innovation économique à la transparence et au renforcement démocratique. Plus concrètement, d'autres objectifs peuvent entrer en jeu comme[1] :

- La promotion de modes de déplacements alternatifs à la voiture
- La conception de services pour faciliter l'accessibilité des personnes handicapées et à mobilité réduite
- Le soutien à un secteur particulier, par exemple celui de la culture
- La sensibilisation d'une partie de la population à telle ou telle question publique

Ces objectifs doivent être clairement définis en amont. En effet des projets open data sont lancés parfois pour des raisons d'affichage, de communication, sans que les buts ne soient clairement posés. Cette étape est nécessaire afin de bien identifier les acteurs de la démarche open data.

b) Les acteurs, avec qui partager les données ?

On distingue trois types d'acteurs dans le jeu de l'open data :

- Les détenteurs de données. Il s'agit souvent d'acteurs publics disposant d'informations qu'ils souhaitent mettre à disposition.
- Les réutilisateurs. Ils utilisent les informations pour proposer des applications, des médiations ou pour leurs besoins spécifiques.
- Les utilisateurs finaux. Ils consultent les données, les visualisent, accèdent aux applications conçues par les réutilisateurs.

Le principe fondateur des projets d'innovation ouverte est de travailler avec des acteurs externes, et notamment avec ceux qui vont réutiliser les données ouvertes. Ces acteurs sont par exemple les acteurs institutionnels, les institutions culturelles, les associations,

[1] CHIGNARD Simon, *Open data : comprendre l'ouverture des données publiques*, Limoges, Editions Fyp, 2012

Droits d'auteur réservés. OU (cc) BY-NC-ND

les communautés, les médias locaux, les citoyens qui réclamaient l'ouverture (comme par exemple la communauté informelle Open Data Toulouse)[1].

c) Les licences, sous quelles conditions partager les données ?

La licence fixe le cadre juridique qui s'applique aux données ouvertes. Elle définit en particulier les droits et les devoirs des réutilisateurs, mais aussi des producteurs. Le choix d'une licence est en réalité un acte politique fort. Ce choix donne une orientation à la démarche et fixe les règles à appliquer, notamment en ce qui concerne :

- Les droits d'utilisation et de modification de la donnée.
- Les droits de réutilisation commerciale et non commerciale.
- Les obligations : l'attribution (mention de la source des données et de la date de dernière mise à jour) ou encore le partage à l'identique.

Fin 2011, les deux licences les plus utilisées en France pour les portails open data sont la licence ouverte (LO) proposée par Etalab[2] et la licence ODbL[3]. La plupart des licences s'accordent *a minima* sur l'obligation d'attribution. Les réutilisateurs doivent impérativement mentionner la source des données qu'ils utilisent. Le droit de modification des données est accordé dans le cas d'ODbL et de la licence ouverte Etalab.

- ODbL est retenue par les initiatives de Nantes, Paris, Grand Toulouse…
- La licence ouverte Etalab évidemment pour le portail gouvernemental data.gouv.fr, mais aussi les villes de Montpellier, Longjumeau…

Concernant l'obligation de partage à l'identique, ODbL impose aux réutilisateurs cette obligation, Etalab non. Celle d'Etalab est jugée plus favorable à la création d'une dynamique économique. Elle est en outre compatible avec plusieurs licences mises en œuvre à l'étranger, notamment celles du gouvernement britannique (Open Government License).

2) Construction des portails

Le portail est la pierre angulaire de la démarche open data. Pour cela, les communautés qui choisissent de mettre en œuvre une démarche open data proposent

[1] Simon Chignard, op.cit.

[2] Cf. la page d'accueil « qui sommes-nous ? » sur etalab.gouv.fr

[3] Cf. résumé de la licence « Open Database License 1.0 » sur vvlibri.org/fr

Droits d'auteur réservés. OU (cc) BY-NC-ND

généralement un sous-domaine data sur le site web de leur organisation (data.nantes.fr par exemple). Ce portail sert notamment[1] :

- À héberger les données sous différents formats.
- En permettre la découverte par les utilisateurs.
- Fournir aux réutilisateurs des informations sur les licences juridiques, les formats, ou encore la documentation des API[2].
- Proposer un espace d'échange et d'interaction avec l'organisation et entre usagers.
- Offrir des informations diverses sur le programme open data : actualités, rencontres et ateliers, concours.
- Mettre en avant les réutilisations les plus exemplaires.

Le portail a donc un rôle technique mais pas seulement, il participe également à l'animation de la démarche et à la promotion des résultats. Pour proposer la découverte des données, on peut proposer plusieurs voies pour y accéder :

- Par la thématique : culture, transport, vie démocratique…
- Par l'émetteur : le service concerné dans le cas d'une collectivité, le ministère pour le portail gouvernemental…
- Par la couverture géographique : un quartier, une commune, un département, une région…
- Par les formats ouverts : CSV, XLS, XML…
- Par la dernière date de MAJ ou de publication
- Par la popularité…

Certains sites proposent un accès par un moteur de recherche. L'internaute tape sa requête (par exemple vélo) et peut ainsi visualiser tous les jeux de données en rapport avec sa demande. Ce mode est bien adapté pour les portails recensant un très grand nombre de fichiers.

[1] CHIGNARD Simon, *Open data : comprendre l'ouverture des données publiques*, Limoges, Editions Fyp, 2012

[2] Source Wikipédia : En informatique, une interface de programmation (abr. API pour Application Programming Interface) est un ensemble normalisé de classes, des méthodes ou des fonctions qui sert de façade par laquelle un logiciel offre des services à d'autres logiciels. Elle est offerte par une bibliothèque logicielle ou un service web, le plus souvent accompagnée d'une description qui spécifie comment ses programmes consommateurs peuvent se servir des fonctionnalités du programme fournisseur.

Droits d'auteur réservés. OU (cc) BY-NC-ND

Si nous devions recenser les caractéristiques désirables d'un portail open data en matière de données, elles seraient les suivantes :

- Des données mises à jour régulièrement :

Le portail doit être en mesure d'afficher la dernière date de mise à jour ainsi que la fréquence de rafraichissement. Par exemple, trop de sources sont annoncées comme annuelles, mais avec une dernière version datant d'il y a deux ou trois ans. Une MAJ régulière permet à l'internaute de savoir que le portail est toujours « vivant », alimenté en données.

- Des données proposées en plusieurs formats :

Si un seul format est proposé, cela signifie que le portail ne fournit en réalité qu'un fichier, et non de la donnée. Ainsi, si seul un fichier Excel est disponible, l'utilisateur peut d'ores et déjà être quasiment certain qu'il aura du nettoyage à réaliser avant d'utiliser le fichier comme une source de données pour d'autres usages. A l'inverse, si le portail fournit également la donnée en CSV, XML, JSON etc, cela veut dire que la donnée est connue de lui sous forme symbolique et que le téléchargement correspondra donc à un export de cette donnée.

- Des données proposées dans un format utilisable :

Comme nous l'avons vu précédemment, le PDF est à proscrire. L'information est certes bien là, mais inexploitable. Autre exemple, si la donnée est géolocalisée, la présence de formats comme SHP ou GeoJSON montre que la donnée a été soigneusement préparée pour des usages cartographiques. En bref il faut privilégier les formats numériques les plus ouverts possibles. La démarche devrait proposer en priorité un format ouvert de type CSV ou XML, ou à défaut, des formats dits standards comme XLS. Utiliser des formats propriétaires est antinomique à la démarche même d'ouverture.

- Des données documentées :

La documentation doit concerner tant la sémantique (que représente chaque champ du fichier ?) que la syntaxe utilisée (le format retenu pour els dates est-il jj/mm/aa, ou aa/mm/jj ?). Pour cela, il est conseillé de maintenir un métafichier qui inclut notamment le

Droits d'auteur réservés. OU (cc) BY-NC-ND

nom et le contenu des jeux de données, le nom du propriétaire et du diffuseur, le périmètre géographique, les dates de publications et de mise à jour etc[1].

- Plusieurs manières d'accéder à la donnée :

Il existe en réalité deux manières d'y avoir accès : soit sous la forme de téléchargement, soit à l'aide d'une interface de programmation (API). Le téléchargement offre l'accès à l'intégralité des données brutes. Mais la mise à disposition sous forme d'API s'adapte mieux à la réutilisation. Elle permet un accès sélectif et s'adapte très bien aux données dont la mise à jour est fréquente. Ainsi, pour les données de type transport et déplacement, ou celles comprenant des horaires d'ouverture, la mise à disposition sous forme d'API doit devenir le standard[2]. Cependant, proposer une API ne dispense pas d'offrir les données en téléchargement, il est même recommandé de proposer cette double option.

En définitive, il y a quantité de manières de publier des données ouvertes. Le choix dépend notamment des compétences maitrisées en interne, de l'environnement technologique de l'organisation, du volume des fichiers et du trafic. Il existe également des solutions clés en main comme CKAN ou encore la plateforme open data d'In Cité Solution (utilisée par Nantes et Rennes) [3].

3) Les jeux de données

Une fois les objectifs définis et le portail mis en place, à quelles données avons-nous accès ?

Nous verrons qu'il existe des approches variées dans la mise à disposition des jeux de données ainsi que des typologies différentes de jeux de données. Enfin, nous nous attacherons à décrire les jeux de données les plus généralement mis en ligne.

a) Les approches

Il y a plusieurs approches possibles concernant les premiers jeux de données, et chacune présente ses propres avantages et limites :

Une approche thématique et sélective : commencer par un jeu de données limité, mais immédiatement réutilisable. Par exemple pour Rennes, le premier jeu de données a

[1] CHIGNARD Simon, *Open data : comprendre l'ouverture des données publiques*, Limoges, Editions Fyp, 2012

[2] Ibid.

[3] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

concerné les vélos en libre-service[1] (emplacement des stations, disponibilité en temps réel…). Il s'agit de données simples à comprendre, incluant des données en temps réel, immédiatement réutilisables. En quelques semaines on a pu constater le développement d'applications mobiles. Puis dans un second temps, l'offre est étoffée par d'autres jeux de données.

Une approche diversifiée : proposer deux ou trois dizaines de jeux de données sur plusieurs thématiques. C'est le choix de Paris, Nantes, Montpellier, Bordeaux ou encore Toulouse. Au lancement de leur portail, entre 10 et 40 données étaient disponibles dans des thématiques variées[2]. Certaines étaient immédiatement consultables, d'autres donnaient lieu à des médiations.

Une approche massive : offrir dès l'ouverture un grand nombre de jeux de données. C'est l'approche de l'État qui a souhaité montrer une implication politique forte dès le départ[3]. Cela a nécessité un travail de préparation en amont important. Au moment de la mise en ligne d'Etalab, l'accent est mis sur la quantité plutôt que sur la qualité. Mais cette approche a engendré des critiques. Depuis, la plateforme a beaucoup évolué et son approche quantitative a été corrigée. La recherche de données est facilitée par un classement plus accessible et le portail est également plus adapté au mode collaboratif autour des données. Le positionnement de l'État est délicat, car il doit assurer un équilibre en fournissant une donnée de niveau national qui est parfois la fusion des niveaux de région ou départements, tout en ne se limitant bien sûr pas à ce rôle d'agrégateur[4]. Et, comme pour tous les portails mais avec un périmètre plus large, il doit en même temps assurer la qualité de la collecte de la donnée, sa pérennité, sa fraicheur etc. Notons que e portail data.gouv.fr est tout de même le point d'entrée principal parmi tous les portails open data lors d'une recherche de données ouvertes.

b) Les typologies

Nous pouvons recenser plusieurs typologies de données ouvertes :

- Leur thématique principale (démographie, urbanisme, transports, économie, vie démocratique, etc…)

[1] GOUIGOUX Jean-Philippe, *Open data : consommation, traitement, analyse et visualisation de la donnée publique*, St Herblain, Editions ENI, 2016

[2] Ibid.

[3] Ibid.

[4] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

- Leur caractère statique ou dynamique (données chaudes ou froides). Les données chaudes sont mises à jour en temps réel. Exemple, la disponibilité des vélos en stations libre-service, ou l'horaire de passage du prochain bus, métro ou tramway. Quant aux données froides, elles sont statiques. Exemple, les noms des rues, mis à jour tous les trimestres, les prénoms les plus populaires tous les ans…
- Leur degré de complexité (permettant ou non une réutilisation simple)
- Leur composante spatiale et/ou temporelle

c) Les principales thématiques

Parmi l'étendu des thématiques proposées par les différents portails open data, voici celles qui reviennent le plus souvent[1] :

- Vie démocratique (résultats électoraux, budgets et finances publiques, attributions de marchés publics…)
- Démographie (état civil, prénoms populaires, population par classe d'âge…)
- Économie (création d'activités économiques, entreprises enregistrées, inscriptions des demandeurs d'emploi, revenus fiscaux…)
- Environnement (qualité de l'air, de l'eau, parcs et jardins…)
- Arts, culture et patrimoine (les équipements culturels, leur fréquentation et tarifs, le patrimoine, plans cadastraux…)
- Urbanisme et habitat (Occupation de l'espace urbain, volumes bâtis et non bâtis, projets d'aménagement urbain…)
- Transports et déplacements (données concernant tous les modes de transports collectifs et individuels, accessibilité aux personnes à mobilité réduite, modes alternatifs…)
- Équipements et services d'intérêt public (liste d'équipements et services, localisation et horaires d'ouverture, statistiques de fréquentation…)
- Localisation et information géographique (données techniques, nomenclature des voies et des adresses, reliefs naturels…)

d) Une absence d'homogénéité

Afin de conclure sur ce sujet, nous pouvons regretter le manque d'homogénéité de ces plateformes open data. En effet, chaque plateforme possède sa propre catégorisation,

[1] CHIGNARD Simon, *Open data : comprendre l'ouverture des données publiques*, Limoges, Editions Fyp, 2012

Droits d'auteur réservés. OU (cc) BY-NC-ND

ce qui ne facilite pas le travail de comparaison. De plus, si les thématiques sont parfois identiques (exemple des vélos en libre-service), les modes de mise à disposition (API, licences, formats) divergent très souvent d'une ville à une autre. Le jeu de données des vélos en libre-service est sous licence ODbL pour data.rennsemetropole.fr alors qu'il est sous licence ouverte sur data.grandlyon.com.

Cette absence d'homogénéité dans les choix et les pratiques est également liée à la genèse de la plupart des projets. Ils ont émergé sur des territoires différents, avec des équipes projets issues de différents métiers (informatique, communication, innovation…). De même, le tissu de réutilisateurs est propre à chaque région. Il faudrait maintenant encourager l'émergence et l'adoption progressive de standards communs pour permettre aux réutilisateurs la possibilité de proposer facilement leurs services d'un territoire à l'autre[1].

B> Les services publics d'archives et le numérique

Le numérique apparait comme un véritable atout pour les services d'archives publiques (AM, AD, AN), c'est pourquoi une majorité d'entre eux se sont lancés dans ce mouvement. Le numérique leur permet en effet de donner à l'utilisateur un accès à distance, de toucher un plus grand public, de valoriser leurs fonds et biens d'autres usages (1).

Pour atteindre tous ces objectifs, les services d'archives doivent mettre l'accent sur le site web en lui-même : ergonomie, facilité de navigation, compréhension du fonctionnement et de la recherche etc (2).

Enfin, pour satisfaire tous les publics, les services d'archives doivent veiller à diversifier les fonds numérisés, afin de refléter le large spectre des fonds conservés. Mais le seul simple accès aux fonds n'est pas suffisant, il faut également mettre à la disposition de l'internaute toute une gamme d'outils lui permettant d'effectuer ses recherches de manière autonome (3).

[1] Simon Chignard, op.cit.

Droits d'auteur réservés. OU (cc) BY-NC-ND

1) Les atouts du numérique pour les services publics d'archives

a) Pourquoi mettre en ligne ?

Aujourd'hui, nous pouvons remarquer que la mise en ligne des archives est devenue une priorité, voire un impératif pour les services d'archives. Mais pour quelles raisons ?

- Il s'agit pour ces services de mettre à disposition des internautes un certain nombre d'archives numérisées ainsi que leurs instruments de recherche correspondants afin de leur permettre de les exploiter au mieux. Pour Frédérique Garnier des AD44, la numérisation des archives était à la base conçue comme support de substitution succédant aux microfilms. Mais finalement la numérisation « est devenue, avec la possibilité technique de diffuser massivement ces images en ligne, un formidable outil de communication[1] ». Cyril Longin, directeur des AM de Saint-Etienne, est du même avis : « La numérisation c'est le même processus que le microfilmage, on sauvegarde l'original. Mais l'intérêt de la numérisation, c'est qu'elle permet la diffusion. Aujourd'hui, je ne connais plus de services d'archives qui continuent de microfilmer. C'est la possibilité de mettre à disposition les formats de substitution[2] ».
- Les archivistes qui ont répondu à l'enquête mettent souvent l'accent sur le fait que l'accès grâce au numérique évite aux personnes de se déplacer. Pour Frédérique Garnier, l'accès « peut se faire sans contrainte horaire et ne nécessite pas de déplacement[3] ». Pour Jeanne Mallet, « les recherches qui nécessitaient autrefois des déplacements sont aujourd'hui très nettement simplifiées[4] ». L'équipe Vitam est du même avis : « L'accès aux archives nécessitait auparavant pour l'usager un ou plusieurs déplacements physiques aux archives ou une demande par correspondance d'un document qui n'aboutissait pas forcément. Avec le numérique, le déplacement physique n'est plus forcément nécessaire ou, du moins, il peut être limité. Les inventaires sont en ligne et permettent de cibler ses recherches et d'optimiser son déplacement dans un service d'archives ».
- Le site internet devient un prolongement des services d'archives, une véritable salle virtuelle. Pour Cyril Longin, le numérique permet de « donner accès à des

[1] Voir annexe n°1

[2] Voir annexe n°5

[3] Ibid.

[4] Voir annexe n°3

Droits d'auteur réservés. OU (cc) BY-NC-ND

documents compliqués à communiquer en salle de lecture, comme les archives orales, les négatifs, les plans « de 17 mètres de long », les plaques de verre, etc [1]».

- La diffusion en ligne permet aussi d'encourager et de promouvoir la recherche notamment sur l'histoire locale.
- La diffusion en ligne permet de satisfaire tous les publics intéressés par les archives, à savoir les généalogistes, les chercheurs, les enseignants et les scolaires, mais cela permet également de toucher un public plus large et diversifié de par la facilité d'accès qu'offre internet. Par exemple, valoriser les fonds conservés par le service d'archives par le biais d'expositions virtuelles est un bon moyen de capter de nouveaux publics. Frédérique Garnier souligne également que cet accès peut toucher « les publics empêchés [2]».
- De manière plus large, cette volonté de diffusion est également à replacer dans un contexte de pression sociétale, qui a besoin de disposer d'information de manière immédiate, libre et sans contraintes. Cela a des conséquences : certains internautes en viennent à douter de l'existence de ce qui n'est pas en ligne.
- Notons également que cela répond à une volonté des gouvernements de prôner le tout numérique et la diffusion d'une grande partie des informations publiques sur internet. Dans un but de transparence et de démocratie, l'Etat cherche ainsi à garantir l'accessibilité citoyenne aux données publiques. Jeanne Mallet le souligne ainsi : « Le numérique permet donc une réelle mise en œuvre du principe selon lequel tout citoyen peut avoir accès aux informations le concernant ou concernant l'État, qui auparavant était plutôt théorique car perçu comme très complexe[3] ».
- C'est aussi une manière de « rajeunir » le métier d'archiviste tout comme le monde des archives. En effet, la première image qui nous vient à l'esprit lorsque l'on parle d'archives, ce sont des vieux papiers poussiéreux au fond d'une armoire dans la cave ou le grenier. La mise en ligne permet de montrer que l'archiviste est toujours en phase avec la société, qu'il souhaite répondre à ses besoins en se plaçant au cœur des nouvelles technologies. Les principes mêmes de l'archivistique se sont adaptés au numérique. Cela nécessite par exemple un compromis entre d'une part le respect de la terminologie archivistique et d'autre

[1] Voir annexe n°5

[2] Voir annexe n°1

[3] Voir annexe n°3

Droits d'auteur réservés. OU (cc) BY-NC-ND

part une vulgarisation par l'usage d'un vocabulaire normalisé et non technique permettant aux néophytes de s'y retrouver.

La mise en ligne d'archives publiques n'est donc pas un « extra » mais bien une dimension actuelle dimension du métier d'archiviste, répondant aux missions de collecte, classement, conservation, communication et valorisation des archives[1].

Le citoyen quant à lieu, qui réclame un accès immédiat et sur l'intégralité des collections, ne voit pas l'envers du décor. En effet, la mise en ligne d'archives publiques demande à ces services beaucoup de temps, une forte implication des agents ainsi que de grands moyens financiers. Encore une fois, cette volonté de diffusion s'inscrit donc bel et bien comme un objectif stratégique dans la politique d'un service.

b) Pour répondre à quels usages ?

La diffusion d'archives en ligne vise à répondre à un certain nombre d'usages :

- En tout premier lieu, cette diffusion doit permettre au chercheur de démarrer son travail et d'effectuer ses recherches à distance. Si celui-ci n'a pas forcément accès à toutes les archives qu'il souhaite consulter, le site doit au moins lui permettre d'avoir accès aux instruments de recherche, de façon à pouvoir préparer sa visite au centre d'archives et ainsi gagner du temps.
- Au-delà de la simple mise à disposition d'archives, le site doit offrir à son public de l'autonomie. En effet contrairement à une vraie salle de lecture où le lecteur est accompagné dans sa recherche par des archivistes, l'internaute est seul devant son ordinateur.
- Enfin, et nous le verrons plus en détails dans le chapitre III, la mise en ligne vise aussi à impliquer l'internaute, dans une démarche de web collaboratif.

c) Quels droits d'accès et d'utilisation ?

Un des objectifs des services d'archives dans cette démarche étant de permettre l'accès citoyen, l'accès à ces archives numérisées est gratuit.

Concernant la réutilisation, notons qu'elle n'est pas libre de droits. En effet les sites web des services d'archives rappellent que la réutilisation des informations publiques est libre, si l'internaute s'engage à ne pas les diffuser publiquement ou à des tiers à des fins

[1] http://ica2012.ica.org/files/pdf/Full%20papers%20upload/ica12Final00407.pdf

Droits d'auteur réservés. OU (cc) BY-NC-ND

commerciales, et si l'information est citée de façon visible et systématique quel que soit son usage, en précisant le lieu de conservation et la cote du document.

La réutilisation d'informations à des fins commerciales est soumise à rétributions, afin de compenser les coûts de mise en œuvre, de numérisation et de mise à disposition sur internet des archives publiques. Rappelons que les services d'archives n'ont pas obligation de mettre en ligne les archives qu'ils conservent, il s'agit d'une démarche proactive de leur part.

2) Les sites web des services publics d'archives

En 2015, le Service interministériel des Archives de France a souhaité mener une enquête[1] sur les sites Internet des archives départementales françaises. En effet, celui-ci souhaite contribuer au développement de la dématérialisation des archives publiques. Cette étude, confiée à un cabinet de consultants, permet de dresser un état des lieux des sites web des AD et de leurs pratiques en ligne.

Pour cela, ce cabinet a mené un examen complet de la totalité des sites des archives départementales (101). Quatre thématiques sont étudiées en profondeur : la stratégie web, la navigation et ergonomie, la disponibilité et la valorisation des ressources archivistiques et enfin les services aux internautes. Nous rapporterons dans ce mémoire les grandes conclusions de ces quatre chapitres, qui nous permettent d'avoir une vue d'ensemble sur cette nouvelle forme d'accès aux archives publiques qu'offre le web.

a) Stratégie web, navigation et ergonomie

Cette étude nous révèle en premier lieu que 75% des sites d'AD sont distincts des sites des Conseils départementaux. Ce chiffre est plutôt satisfaisant, car la mission des AD est si particulière qu'elle mérite de posséder son propre site internet.

Lorsque les AD sont représentées via le site du Conseil départemental, il s'agit plutôt d'un site d'information, qui va donner par exemple l'adresse, les horaires d'ouverture, les missions générales mais rien de plus. Ce qui est assez frustrant pour le chercheur, qui n'y trouvera aucun fonds numérisé. Aujourd'hui en 2017, ces chiffres ont déjà dû évoluer faisant réduire la proportion de l'hébergement sur les sites des conseils départementaux.

[1] Portail France archives : http://www.archivesdefrance.culture.gouv.fr/archives-publiques/chiffres-clefs-rapports-et-etudes/etudes-publiees/etudes-sur-les-sites-internet/

Droits d'auteur réservés. OU (cc) BY-NC-ND

Sur un site propre aux archives départementales, l'internaute fait avant tout des recherches. En termes de stratégie, le cabinet nous indique que la très grande majorité des AD proposent sur leur site à la fois des contenus éditoriaux et des ressources numériques (95%). Un tiers de ces sites vont même au-delà, en proposant des fonctions collaboratives ou contributives. Cette démarche donne immédiatement une image moderne de l'activité des AD, tournée vers les nouvelles technologies. La moitié des sites étudiés se présentent comme modernes d'un point de vue graphique. Cela montre bel et bien un effort de la part des AD, qui souhaitent valoriser leur activité à travers un design travaillé. De plus, un site sur deux est mis à jour de manière fréquente et valorisée.

Concernant l'ergonomie, celle-ci est très variable d'un site à l'autre. Elle diffère notamment en fonction de l'outil utilisé ainsi que du contenu hébergé par le site. Certains sites ont une ergonomie très simple, qui peut être tout à fait satisfaisante si le site n'héberge que peu de contenu ou de ressources. Cependant, dans le cas où le site propose un grand nombre de ressources et de contenus, une ergonomie trop simpliste peut s'avérer être une faiblesse. L'internaute risque de ne plus retourner sur le site.

Deux tiers des sites offrent une navigation dite « facile » voire « très facile » selon les critères du cabinet de consultants. Ce chiffre est satisfaisant. Cette donnée est importante, car comme nous le verrons par la suite, les archives numérisées présentes sur les sites d'AD sont d'abord liées à la généalogie. Or il se trouve que le public intéressé par ce domaine (hors professionnel) est majoritairement retraité. Un accès facile doit être offert à ce public afin qu'ils ne délaissent pas cette méthode de recherche et continuent de fréquenter le site internet des AD.

Enfin, en termes de *responsive design*[1], en 2015 seulement 20 sites sur 101 sont *mobile friendly*. Aujourd'hui en 2017 il est fort probable que ce chiffre ait considérablement augmenté.

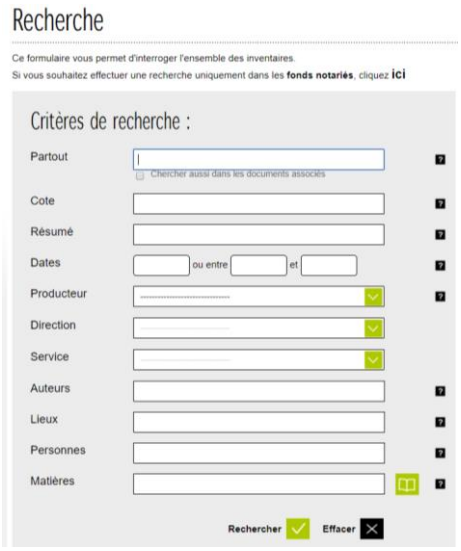
b) Moteur de recherche

Nous pouvons remarquer que les moteurs de recherche dans les fonds sont systématiquement présents. Certains proposent même des critères de tri.

[1] Le responsive design est l'adaptation du site selon l'appareil depuis lequel il est lancé. Notamment lors de la visite du site via un smartphone. Bouton spécifique de menu, menu redimensionné, largeur d'écran respectée…

Droits d'auteur réservés. OU (cc) BY-NC-ND

Ci-dessous une capture d'écran[1] du moteur de recherche de l'inventaire des archives des AD de Loire-Atlantique. Celui-ci propose toute une sélection de critère : producteur, direction, service, auteur, lieu etc.



c) Portails et moissonnage des données

Début 2017, et sur initiative du SIAF, un nouveau site est apparu, appelé FranceArchives.fr. Il s'agit en réalité d'un portail, qui permet de trouver les références de plusieurs millions de documents d'archives conservés dans les services publics d'archives. Ce portail dispose d'un moteur de recherche qui interroge de manière simultanée les inventaires de toutes les institutions partenaires du site en France. Comme le projet est assez récent, peu d'institutions sont aujourd'hui représentées, mais le site a vocation à s'étendre dans les mois et les années à venir.

Pour Hervé Lemoine, directeur du SIAF, « l'objectif de FranceArchives est de révéler ces fonds dispersés sur le territoire national, et encore parfois méconnus. Il devient en effet indispensable de diffuser les documents sur internet de façon techniquement simple et scientifiquement cohérente à l'échelle nationale ». [2]

Le directeur du SIAF rajoute que ce site doit « répondre aux impératifs de l'indexation orchestrés par les grands moteurs de recherche de type Google, Bing ou Qwant ».[3] De

[1] https://archives.loire-atlantique.fr/jcms/chercher/inventaires-des-archives/recherche-fr-p1_8558

[2] Lettre du COEPIA (conseil d'orientation de l'édition publique et de l'information administrative, entretien avec Hervé Lemoine : http://www.gouvernement.fr/sites/default/files/contenu/piece-jointe/2016/12/lettre_du_coepia_ndeg_40.pdf

[3] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

plus, ce site a pour vocation à « aider les services d'archives à développer et à améliorer leur politique de diffusion numérique, y compris grâce aux médias sociaux ».[1]

Concernant les autres bases de données que présente le web, notons qu'en 2015 selon l'étude du cabinet de consultants, 5% des sites d'AD avaient un lien vers une base de données externe, et 6% participaient à un portail régional. Un bon exemple est la banque numérique du savoir d'aquitaine (les AD24, AD33, AD47 y participent). Au-delà des portails régionaux, certains sites renvoient vers d'autres portails ou ressources : par exemple, les AD23 renvoient vers le site « mémoire des hommes », ou encore les AD44 qui renvoient vers le site de la cinémathèque de Bretagne.

En définitive, si nous devions résumer les facteurs clés de succès d'un site web d'AD, ils seraient les suivants :

- Ergonomie réfléchie et simplifiée, intuitivité
- Site conforme aux usages du Web (*User Friendly*)
- Des ressources complètes, riches et variées
- Les meilleurs moyens de les exploiter…
- … notamment par un moteur de recherche puissant et unifié
- Référencement optimisé
- Lieu de sérendipité ?

d) Quelques données chiffrées

Chaque année, le SIAF réalise des statistiques[2] auprès des services publics d'archives au titre du contrôle scientifique et technique (les chiffres annoncés ci-dessus proviennent du rapport le plus récent disponible sur FranceArchives.fr, à savoir 2015). Les indicateurs demandés par le ministère auprès de ces institutions concernent aussi bien le fonctionnement des services que leur action culturelle, la collecte, le classement des archives ou encore la numérisation des archives. Concernant notre sujet, trois thèmes nous intéressent dans ces chiffres : numérisation, mise en ligne et consultation en ligne.

Voici les chiffres les plus significatifs correspondants à ces trois thèmes.

Numérisation :

[1] Lettre du COEPIA, op.cit.

[2] Portail France Archives : https://francearchives.fr/article/37978

Droits d'auteur réservés. OU (cc) BY-NC-ND

- Nombre total de pages numérisées en 2015 : 22 millions.
- Nombre total de pages numérisées : 433 millions, dont état civil : 247 millions. Cela signifie qu'en 2015, 57% du nombre total de pages numérisées provient de l'état civil.
- Nombre total d'images numérisées : 12 millions. Dont 460 000 images pour le cadastre. Soit moins de 4%.

Mise en ligne :

- Le premier site a été lancé en 2000, celui des Côtes d'Armor.
- Moyenne du nombre de pages mises en ligne par AD : 3,5 millions, dont 2,4 millions d'état civil, soit environ 70%.
- Total : 347 millions de pages en ligne (335 millions en 2014)
- Total : 8 millions d'images mises en ligne

Consultation en ligne :

- Plus de 2 milliards de pages et d'images vues (2 252 094 421 pages).
- 42 millions de connexions
- 11 millions de visiteurs uniques (en moyenne 134 000 par AD)

Certains chiffres sont à prendre avec des pincettes : le Loir-et-Cher (41) enregistre 415 599 visiteurs uniques alors que le département fait 333 000 habitants… Le site des AD du Loir-et-Cher n'est certes pas réservé à ses uniques habitants, mais on se doute qu'il y a un problème. Cela montre bel et bien qu'il est très difficile de mesurer la réelle ampleur de la consultation en ligne. Cependant, l'ampleur des chiffres nous montre que la numérisation des archives et leur mise en ligne est un véritable enjeu pour les services publics d'archives, et que ce chiffre ne cesse de grimper d'années en années.

3) Les archives numérisées et les outils disponibles en ligne

L'accès aux archives en ligne s'appréhende à travers deux axes : d'une part la numérisation de ces archives, et d'autre part le signalement de ces archives numérisées, via la diffusion d'inventaires en ligne.

Dans cette partie nous allons analyser l'offre de service des sites du point de vue de l'ensemble des ressources archivistiques mises à disposition des internautes sous leur forme numérique, à savoir principalement : les fonds numérisés et les instruments de recherche.

Droits d'auteur réservés. OU (cc) BY-NC-ND

Nous traiterons dans un premier lieu de la richesse et la diversité des ressources mises à disposition d'une part, avant d'étudier les conditions d'accès et de recherche proposées aux internautes sur ces ressources.

a) Richesse et diversité des fonds numérisés

La question que nous nous posons est la suivante : quelles sont les archives numérisées disponibles en ligne ?

Voici le classement des fonds majoritairement présents sur les sites d'AD[1] :

- En tête, nous trouvons l'état civil et tables décennales : présents sur 89% des sites d'AD.
- Puis ce sont les plans cadastraux : présents sur 82% des sites d'AD.
- Les registres paroissiaux arrivent en troisième position : ils sont présents sur 80% des sites d'AD.
- Ensuite ce sont les registres matricules militaires : présents à 75% sur les sites d'AD.
- Le recensement des populations est présent sur 66% des sites d'AD.
- Les fonds iconographiques sur quant à eux présents sur 65% des sites d'AD.

Il s'agit principalement de fonds sériels, en lien avec les attentes des publics généalogistes. La presse arrive loin derrière avec seulement 30%. Mais le cabinet indique que plus de la moitié des sites étudiés proposent des fonds complémentaires, très variables. On peut par exemple trouver :

- Les répertoires d'écrou
- Les cahiers de doléances
- Les fonds associés à la Première Guerre mondiale (notamment dans le cadre du 100ème anniversaire)

Sans oublier des fonds atypiques comme les herbiers, les armoiries, les buvards publicitaires, etc. A l'avenir, certains fonds pourraient devenir des items à part entière, comme les cahiers de doléance par exemple.

[1] Portail France archives : http://www.archivesdefrance.culture.gouv.fr/archives-publiques/chiffres-clefs-rapports-et-etudes/etudes-publiees/etudes-sur-les-sites-internet/

Droits d'auteur réservés. OU (cc) BY-NC-ND

Ces fonds complémentaires sont souvent très riches, et mettent en valeur l'histoire ainsi que les spécificités du territoire concerné (dans le cadre de l'histoire régionale).

- En lien avec la géographie : inscriptions maritimes, matricules de navire, journaux et rôles de bord, rapport de mer des capitaines pour les AD44[1] ou encore le fleuve Charente[2] pour les AD16.
- En lien avec l'histoire : archives militaire de la guerre de Vendée[3] pour les AD85, « un jour en 44[4] » des AD de la Manche.

D'un point de vue technique, notons que la numérisation de ces archives est généralement réalisée entre 300 et 400 dpi de façon à conserver une bonne qualité d'image. En effet pour certains documents, comme par exemple l'état civil, le chercheur peut avoir besoin de recourir au zoom afin de déchiffrer les écritures.

Enfin, des craintes peuvent être relevée concernant un effet pervers des archives disponibles en ligne. En effet, pour Frédérique Garnier, « si la numérisation met en lumière quelques fonds, souvent déjà les plus consultés, toutes les archives ne peuvent être numérisées et la plus grande partie des documents restent et resteront dans l'ombre[5] ». Autre crainte pour l'équipe Vitam, étant donné que « l'usager n'a accès en ligne qu'à certains corpus (tout n'est pas accessible), il peut manquer de visibilité sur des fonds complémentaires, non visibles et accessibles en ligne. Ce qui peut finalement le faire passer à côté de certains types d'archives[6] ».

b) Conditions d'accès et de recherche

La première manière d'accéder à des archives numérisées est l'approche par fonds (56% des sites d'AD proposent cet accès par fonds selon l'étude[7]). Mais il existe aussi une approche mixte. Cette approche propose généralement un accès par fonds ainsi qu'un accès par thèmes. Voici quelques exemples :

[1] Site web des AD44 : https://archives.loire-atlantique.fr/jcms/chercher/archives-numerisees/marine/rapports-de-mer-des-capitaines/rapports-des-capitaines-a-l-amiraute-de-nantes-fr-t1_6165?accepte=true&portal=c_5110

[2] Site web des AD16 : http://archives.lacharente.fr/arkotheque/consult_fonds/index.php?ref_fonds=3

[3] Site web des AD85 : http://www.archives.vendee.fr/Consulter/Archives-numerisees

[4] Site web des AD50 : http://www.archives-manche.fr/blog/t1448

[5] Voir annexe n°1

[6] Voir annexe n°4

[7] Portail France archives : http://www.archivesdefrance.culture.gouv.fr/archives-publiques/chiffres-clefs-rapports-et-etudes/etudes-publiees/etudes-sur-les-sites-internet/

Droits d'auteur réservés. OU (cc) BY-NC-ND

- Approche par grands types de fonds : état civil, registre paroissial, iconographie, presse…
- Approche par sujet : Première Guerre mondiale
- Approche conceptuelle : trésor des chartes des ducs de Bretagne (AD44), Commission des Monuments historiques du département de la Gironde (AD33)…

Certains thèmes d'actualité, comme le centenaire de la Première Guerre mondiale, constituent parfois un portail à part entière comme aux AD28. On y retrouve bien sûr des lettres de poilus, mais aussi des photographies, des témoignages…

Toujours selon le cabinet d'études, 70% des sites proposent l'approche mixte.

Comme nous l'avons déjà signalé, permettre l'accès aux différents fonds n'est pas suffisant du point de vue de l'archiviste. Tout comme dans une salle de lecture ou le lecteur est accompagné, le site internet, ou salle de lecture virtuelle, tend aussi à accompagner l'internaute dans son cheminement. Pour cela, les sites des AD proposent très souvent de l'aide à la recherche. Selon l'étude du cabinet de consultants[1], 75% des sites des AD proposent des fiches d'aide d'une à cinq pages, et 25% d'entre eux proposent des guides de recherches supérieurs à 5 pages.

En plus de cette aide à la recherche, le chercheur a besoin d'avoir accès au cadre de classement ainsi qu'aux instruments de recherche du service public d'archives. Notons qu'un cadre de classement est disponible dans près de 80% des cas. Mais, comme nous l'avons signalé plus tôt, certaines AD ont choisi délibérément de ne pas le présenter pour s'orienter dans une démarche tournée vers le grand public.

Concernant les fonds, la très grande majorité des sites d'AD donnent accès à l'état de leurs fonds. En consultant cet état des fonds, le chercheur peut savoir d'emblée quelles sont les archives en la possession des AD. Mais la profondeur de recherche dans le fonds varie beaucoup d'un site à l'autre, ce qu'on peut regretter. Les instruments de recherche sont associés aux séries du cadre de classement dans 72% des cas, mais avec également un niveau de granularité qui diffère. Lorsqu'il y a une forte profondeur, le chercheur peut

[1] Portail France archives : http://www.archivesdefrance.culture.gouv.fr/archives-publiques/chiffres-clefs-rapports-et-etudes/etudes-publiees/etudes-sur-les-sites-internet/

Droits d'auteur réservés. OU (cc) BY-NC-ND

ainsi gagner beaucoup de temps en programmant sa visite aux AD, car il aura noté toutes les cotes qui l'intéressent.

De même, les sites qui proposent des instruments de recherche dynamiques (40%) à partir du cadre de classement permettent un gain d'efficacité et de lisibilité.

L'étude nous révèle également que près de la moitié des sites d'AD intègrent des instruments de recherche au format EAD ou XML. Ces formats améliorent l'ergonomie des ressources archivistiques et facilite leur diffusion et leur intégration dans les portails nationaux (FranceArchives) ou européens (APEx : Archive portal Europe. Il s'agit d'un site de mise en commun d'instruments de recherche produits par les services d'archives des pays membres de l'UE). De plus, quand les instruments de recherche sont disponibles au format EAD, les fonds numériques associés sont directement accessibles depuis ces instruments.

La conclusion de l'étude menée par le cabinet de consultants permet de classer les sites d'AD en trois catégories :

- Environ la moitié des sites se positionnent de manière satisfaisante voire très satisfaisante tant par rapport à la richesse et à la diversité des fonds que par le niveau d'accès et de recherche dans les ressources.
- Une trentaine de sites ne proposent que peu de ressources et des fonctions de recherche limitées.
- Entre ces deux premières catégories, se placent quelques sites qui soit disposent de bonnes fonctionnalités mais de peu de ressources, soit de fonds riches et variés mais s'appuyant sur des conditions d'accès et de recherche limitées.

Le constat est donc mitigé, la marge de progression demeure importante. Certains sites présentent un réel potentiel d'évolution, en effet on observe que les fonctions liées aux ressources archivistiques sont généralement plus développées que celles liées au web. On constate donc que c'est bien l'archiviste qui est au centre de cette démarche.

Cette analyse a permis de constater qu'il existe une grande diversité dans les moyens déployés pour la mise en œuvre des sites web, ainsi que des stratégies web différentes. Certaines AD placent leur site web au cœur de leurs actions, en proposant un site récent,

Droits d'auteur réservés. OU (cc) BY-NC-ND

ergonomique, offrant une image moderne, attractive, ouverte et dynamique du métier des archives.

De manière générale, nous pouvons constater une véritable volonté d'utiliser internet et les nouvelles technologies pour valoriser l'activité des AD au-delà des salles de lecture et des expositions. L'internaute a ainsi accès, depuis chez lui, à une masse considérable de fonds d'archives, d'outils, de fonctionnalités qui représentent un fort potentiel.

Enfin, nous pouvons nous demander si les sites web des services publics d'archives constituent une fin en soi ?

En définitive ceux-ci ne tendent pas et ne tendront jamais à numériser et mettre en ligne toutes les archives qu'ils conservent. Le site internet vise davantage à offrir de la visibilité, à faciliter l'accès aux instruments de recherche ainsi qu'aux documents d'archives régulièrement sollicités.

Pour les archivistes interrogés, l'accès grâce au numérique apporte également son lot d'inconvénients.

Pour Jeanne Mallet, le numérique un impact sur la façon dont les usagers perçoivent l'accès : « Dans la mesure où il est aujourd'hui possible de transmettre "instantanément" et "sans frais" les documents à l'autre bout de la planète, les personnes qui souhaitent avoir accès aux documents comprennent mal les obligations qui leur sont parfois faite de se déplacer, voire de payer les reproductions des documents qu'ils demandent lorsqu'ils sont numérisés pour permettre leur envoi[1] ». Cyril Longin est du même avis : « L'inconvénient en termes d'accès, c'est que les gens ne se rendent pas compte du travail en amont[2] ».

Le numérique peut donc également avoir un aspect négatif, lié surtout à une méconnaissance des enjeux notamment financiers qu'il représente pour les services : « l'inconvénient est une mauvaise compréhension de la façon dont le numérique fonctionne, notamment la croyance paradoxale en la gratuité du numérique (alors que les campagnes de numérisation se calculent en millions d'euros) de la part du public. Cette

[1] Voir annexe n°3

[2] Voir annexe n°5

Droits d'auteur réservés. OU (cc) BY-NC-ND

croyance se trouvent également chez les décideurs, qui ne prennent pas forcément en compte les coûts de conservation et de communication qui sont pourtant bien réels[1] ».

De plus il ne faut pas tomber dans le piège de croire que tout ce qui n'est pas sur internet n'existe pas. Les contenus archivistiques en ligne ne constituent qu'une manière d'accéder à l'information. De plus pour Frédérique Garnier, « l'habitude d'accéder à des documents numérisés engendre parfois une frustration chez les usagers lorsque les archives ne le sont pas ou que la consultation des inventaires ne débouche pas sur une communication immédiate des documents[2] ».

Frédérique Garnier rajoute que cet accès est tributaire de la technique et peut engendrer des problèmes : « si le site internet est hors service, le service est interrompu[3] ». De plus, il engendre de nouvelles préoccupations pour le service public d'archives : « pérennité des formats, versions de navigateurs, anti-virus, outils de visualisation, stockage, environnement du poste de travail… [4]». Jeanne Mallet rajoute que « la conservation des documents nativement numériques pose nombre de questions[5] ».

A travers cette partie, nous avons vu que les administrations doivent répondre à des obligations de publication en ligne d'un certain nombre de données publiques. Les administrations répondent à cette obligation par le biais de portails open data, et les services publics d'archives par le biais de sites internet qui met en valeur le patrimoine qu'ils conservent dans leurs murs. Mais sont-ils toujours aujourd'hui les seuls acteurs dans le domaine de l'accès aux archives publiques ?

III> VERS DE NOUVEAUX ACTEURS ?

Les services publics d'archives, bien qu'ils aient pris le train en marche de la diffusion des données et de leurs archives numérisées en ligne, ont tout de même perdu leur monopole de diffusion. En effet, de nouveaux acteurs sont apparus, la donnée publique étant devenue un véritable « filon ».

[1] Ibid.

[2] Voir annexe n°1

[3] Ibid.

[4] Voir annexe n°1

[5] Voir annexe n°3

Droits d'auteur réservés. OU (cc) BY-NC-ND

Nous illustrerons ce constat à travers deux cas, celui d'un EPIC à savoir la SNCF (A), et celui d'une entreprise privée, nommée Filae (B). Nous aurions pu étudier de nombreux autres cas mais nous avons décidé de nous focaliser sur ceux-ci car ils nous semblent représentatifs de l'émergence de ces nouveaux acteurs, et de l'intérêt suscité par les données publiques.

A> Le cas d'un EPIC : la SNCF

Le groupe SNCF est composé de trois EPIC : SNCF, SNCF Réseau et SNCF Mobilités. En 2014, le groupe décide de se lancer dans une démarche de transparence et ouvre une partie de ses données. Pour le groupe, ouvrir les données ne peut avoir qu'un impact positif. L'open data est vu comme un « accélérateur d'innovation au service de la mobilité de tous »[1].

En effet, la SNCF transporte chaque jour des millions de voyageurs. Afin de mener à bien cette mission, l'entreprise a besoin de créer et de travailler avec un grand nombre de données. Ces données concernent les horaires des trains, les équipements dont disposent les gares, la régularité des trains, l'accessibilité en gare, etc. En cela, l'Open Data ainsi que les API que propose la SNCF sur son site data.sncf.com sont pour l'entreprise un moyen de favoriser l'innovation. De plus, cette démarche participe à la transparence de la SNCF, qui par ce biais délivre un grand nombre d'informations à ses usagers.

Au-delà du simple usager, c'est aussi l'entreprise qui est visée par cette démarche. Des développeurs, des chercheurs, des entreprises, des *data scientists* peuvent s'approprier ces données afin de les réutiliser. Nous parlerons plus en détails de la réutilisation des données dans le troisième chapitre. Grâce à ces données, ils pourront développer des applications visant par exemple à optimiser et valoriser le temps de voyage, ou encore à fournir de nouveaux services innovants aux voyageurs.

La SNCF a mis en place deux plateformes dédiées à cette démarche : une première plateforme pour ses jeux de données, et une deuxième pour ses archives numérisées. Ces deux plateformes sont le fruit d'un travail collaboratif entre archivistes du groupe, informaticiens et juristes[2]. Lors du lancement de sa plateforme open data intitulée data.sncf.com, 18 jeux de données étaient alors disponibles[3]. Aujourd'hui, 204 jeux de

[1] https://data.sncf.com/about

[2] Bruno Texier, « SNCF : l'open data est sur de bons rails », Archimag n°303, avril 2017.

[3] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

données sont accessibles sur son site[1]. Le plus gros des jeux de données sont arrivées en 2017 (137). Le site en lui-même est très ergonomique, facile à appréhender et utiliser. Différentes « vues » sont proposées afin de représenter l'information : analyse, carte, image, vue personnalisée. Les producteurs des données sont également bien dissociés : SNCF Réseau, Transilien, SNCF Voyages, Gares & Connections etc.

Par rapport aux jeux de données en eux-mêmes, le site présente de nombreux thèmes. Par ordre croissant en nombre de jeux de données nous avons : rapport financier (53), description du réseau (38), gares et infrastructures (22), social et ressources humaines (20), informations institutionnelles et financières (15), services voyageurs (13), horaires et itinéraires (10), etc.

Concrètement, le site permet d'avoir accès au trafic des trains en temps réel, au recensement des incidents, mais aussi des objets trouvés… Pour la SNCF, cette démarche d'ouverture a également pour but de montrer que ses trains n'arrivent pas toujours en retard, et donc de rehausser son image. En effet, grâce aux chiffres de la régularité, on apprend par exemple qu'en décembre 2016, 9 Ter sont arrivés à l'heure pour 1 en retard[2]. Mais c'est à double tranchant, car par exemple pour les TGV, les chiffres sont nettement moins bons. En effet, en juin 2017, 4 TGV sont arrivés à l'heure pour 1 en retard[3].

Cette régularité peut être vérifiée de manière très précise, par exemple pour un trajet en particulier. L'exemple ci-dessous nous montre la régularité mensuelle des TGV pour la ligne Brest-Paris Montparnasse sur l'année 2016. En novembre 2016, la régularité était de 86.1%. (31 trains en retard sur 223 ayant circulé)[4].

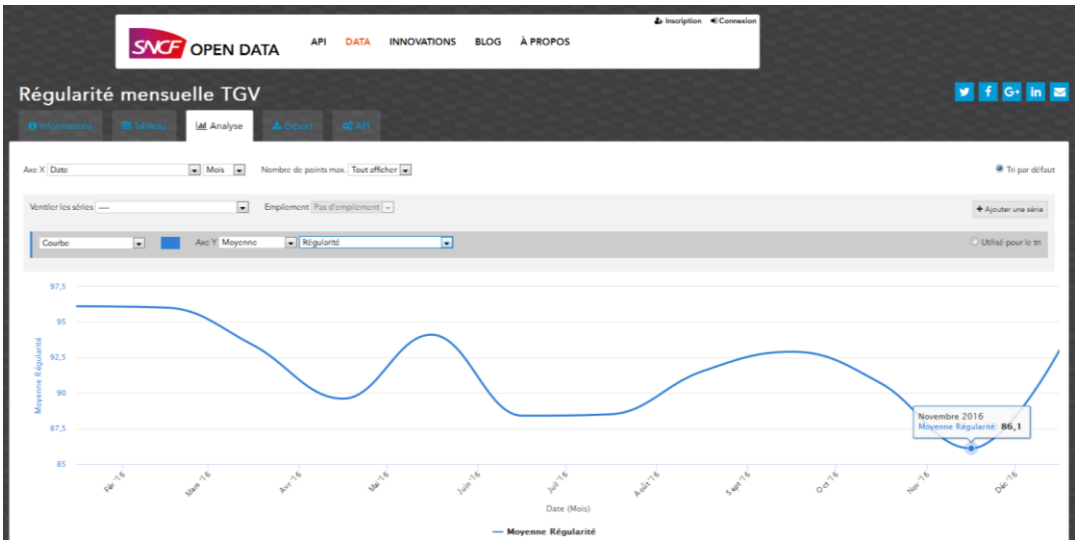
[1] Portail open data de la SNCF https://data.sncf.com/ consulté le 08/06/17

[2] http://www.sncf.com/fr/transparence/barometre

[3] http://www.sncf.com/fr/transparence/barometre

[4] Portail open data de la SNCF : https://data.sncf.com/ consulté le 13/04/17

Droits d'auteur réservés. OU (cc) BY-NC-ND



D'autres jeux de données sont pour le moins surprenants, notamment celui consacré aux objets trouvés dans les gares en temps réel. Ainsi, nous apprenons que le 13/04/17, un doudou a été retrouvé à 10h dans la gare de Brest.

Enfin, concernant les formats, tout le catalogue est disponible aux formats suivants : XLS, CSV, RDF, flux RSS. La variété des formats proposés permet une réutilisation facilitée.

Précisons que pour protéger ses données, la SNCF a mis en place une licence spécifique qui détermine les droits et les devoirs des réutilisateurs. Il s'agit d'une licence Creative Commons de partage des conditions initiales à l'identique (ou *share-alike*). De plus, la SNCF met à la disposition des réutilisateurs une nouvelle API qui fournit des données en temps réel. Quatre fonctionnalités peuvent ainsi être interrogées grâce à celle-ci :

- Itinéraires en train
- Arrêt en gare
- Grilles horaires de ligne
- Recherche automatique des gares

Pour le directeur du programme Transparence et open data de la SNCF, Guillaume Leborgne, ce choix d'ouvrir les données en 2014 a donné de l'avance à la SNCF[1]. Pour Romain Lalanne, responsable Open data à la SNCF, la démarche proactive de la SNCF d'ouvrir ses données vise deux objectifs : « d'une part, améliorer la diffusion de l'information voyageurs surtout pour les transports quotidiens, sur tous les canaux

[1] Bruno Texier, op.cit.

Droits d'auteur réservés. OU (cc) BY-NC-ND

possibles, avec l'idée qu'il est possible de personnaliser plus les informations et en s'appuyant sur les start-ups ; et d'autre part, participer à la transformation de l'entreprise autour de la donnée[1] ».

En plus de ce portail open data, la SNCF s'est également dotée d'une plateforme permettant la consultation d'un grand nombre d'archives en ligne. Le fonds d'archives de la compagnie est très riche. Celui-ci est constitué entre autres d'environ 500 000 photographies et de plus de 5000 films[2]. La SNCF lance cette plateforme en ligne intitulée open archives[3] en 2016, soit deux ans après sa démarche d'open data. Cette plateforme rassemble plus de 1500 documents consultables en ligne. Ce volume d'archives a vocation à s'accroitre au fil du temps.

Les archivistes de l'entreprise ont été les promoteurs de ce projet. En effet, les archivistes y ont trouvé leur place, « car la SNCF a une très forte culture du document » selon Guillaume Leborgne. Agnès d'Angio-Barros, directrice du Sardo, le service archives et documentation de la SNCF, rajoute que « les archivistes ont un rôle clé dans la mise en avant des données ». En effet, ils sont essentiels au moins à deux sujets, « la connaissance des fonds d'archives et la connaissance du droit des archives[4] ».

Sur le site Open Archives, la SNCF justifie cette démarche de mise en ligne d'un certain nombre de ses archives de la manière suivante[5]:

> Les archives SNCF sont des archives publiques et à ce titre elles peuvent être, suivant certaines règles, consultées par tous. Pour rendre plus aisée cette consultation jusque lors uniquement accessible sur des lieux physiques, ce site web a pour vocation de présenter une sélection de documents, films et photographies, conservés dans les centres du Service Archives Documentation de SNCF (Centre national des archives historiques, au Mans ; Centre national des archives du personnel, à Béziers ; Médiathèque).

[1] Ibid.

[2] Ibid.

[3] http://www.sncf.com/fr/transparence/open-archives

[4] Bruno Texier, op.cit.

[5] http://openarchives.sncf.com/pourquoi-ouvrir-nos-archives

Droits d'auteur réservés. OU (cc) BY-NC-ND

C'est donc la volonté d'élargir l'accès qui est au cœur de ce projet.

Ces documents présentent eux-aussi une grande variété : photographies, plans de gare, documentation, films… Et tous ces documents sont accessibles via une interface qui propose plusieurs modes de recherche : mot-clef, thème, localisation, période…

Citons quelques thèmes mis en avant par le portail open archives : la construction des premiers chemins de fer français, la reconstruction du réseau ferré après la Seconde Guerre mondiale, l'électrification des lignes, la grande vitesse, le tourisme ferroviaire, la politique commerciale de la SNCF, l'architecture des gares etc.

Concernant la rechercher par mots-clefs, voici un exemple de choix que le site propose[1] :

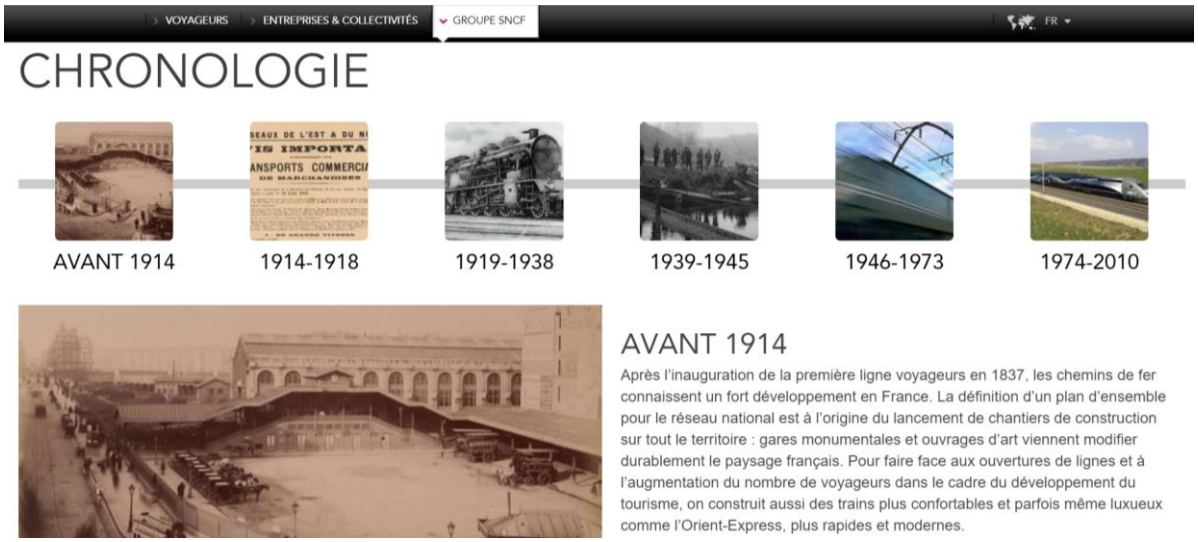
Avec la recherche par localisation, une carte s'affiche à l'écran, et l'internaute n'a plus qu'à cliquer sur la zone qui l'intéresse[2] :

[1] Site web Open archives : http://www.sncf.com/fr/transparence/open-archives

[2] Site web open archives : http://www.sncf.com/fr/transparence/open-archives

Droits d'auteur réservés. OU (cc) BY-NC-ND

Enfin, la recherche par période offre à l'internaute une frise chronologique. Puis, en cliquant sur la période qui l'intéresse, un grand nombre de ressources liées à cette période s'affichent[1] :



Ce portail Open Archives est un moyen pour la SNCF de gagner en visibilité. Pour Agnès d'Angio-Barros, ce portail « est aussi une vitrine derrière laquelle se cache un travail sérieux et règlementairement bien appuyé[2] ».

[1] Ibid.

[2] Bruno Texier, op.cit.

Droits d'auteur réservés. OU (cc) BY-NC-ND

Cette démarche d'ouverture lancée par la SNCF est un succès pour l'entreprise. La compagnie ferroviaire pose désormais comme principe l'ouverture des données, la confidentialité devenant l'exception.

B> Le cas d'une entreprise privée : Filae

Créée en 1994, la société Filae, autrefois généalogie.com, propose aux internautes le moyen de retrouver leurs ancêtres grâce à une base de données qu'elle s'est constituée via la numérisation de l'état civil français. Elle souhaite en cela « faciliter l'accès du grand public à ses racines[1] ». Ce service est basé sur une plateforme Big data associée à des algorithmes utilisant le « *machine learning*[2] ». Par le biais d'un moteur de recherche unique, les internautes peuvent construire leur arbre généalogique à partir d'archives numérisées, qui ont été transcrites et indexées.

Or, en décembre 2016 a lieu « l'affaire Filae ». En effet, cette entreprise a profité de l'absence de mise en conformité des archives départementales avec la loi Valter pour capter et aspirer leurs archives numérisées et les réutiliser à des fins commerciales. Ce qui provoque un véritable tôlé dans le milieu des archivistes français.

Comment Filae a-t-elle pu s'approprier la quasi-totalité de l'état-civil français (environ 100 millions d'images numérisées) ?

Comme nous l'avons vu précédemment, la loi Lemaire institue un principe d'Open Data par défaut. Mais un secteur de la culture et du patrimoine a conservé une position de privilège pour ne pas participer à cette dynamique d'ouverture. C'est notamment l'effet de la loi Valter. Ce texte et son décret d'application permettent aux établissements culturels (bibliothèques, archives, musées) – là où la quasi-intégralité des autres administrations sont désormais soumises à un principe de gratuité – de continuer à lever des redevances pour la réutilisation d'informations issues des opérations de numérisation de leurs collections (et des métadonnées associées).

Or il se trouve que la loi Valter comporterait une faille[3]. En effet, la majorité des services d'archives départementales avaient mis en place des tarifs de réutilisation au titre de la numérisation de leurs collections. Ces tarifs étaient établis sur le fondement de la loi du

[1] Site web Filae : http://www.filae.com/ressources/qui-sommes-nous/

[2] Apprentissage automatique

[3] Lionel Morel : https://scinfolex.com/2016/12/13/une-enorme-faille-dans-la-loi-valter-sur-les-donnees-culturelles/

Droits d'auteur réservés. OU (cc) BY-NC-ND

17 juillet 1978 relative à la réutilisation des informations publiques. Mais l'article 10 de la loi Valter prévoyait une période transitoire au cours de laquelle les établissements devaient mettre en conformité leurs licences avec la loi, notamment en ce qui concerne les modalités de calcul des redevances qui ont été modifiées. Mais les départements n'ont pas respecté ces délais pour mettre à jour leurs licences et règlements, ce qui les a rendus caducs. C'est à ce moment que Filae en a profité pour aspirer toutes ces données. L'entreprise est aujourd'hui le seul acteur à disposer de la quasi-intégralité de l'état civil en un point centralisé.

Sur son blog, Filae.com soutient que les Lois Valter et Lemaire ont eu pour effet de faire passer les données d'archives en Open Data[1] :

> « Ce projet a été rendu possible grâce aux travaux de numérisation des registres originaux principalement réalisés par les départements français. En vertu de la loi Valter « relative à la gratuité et aux modalités de la réutilisation des informations du secteur public » et de la loi Lemaire « pour une République Numérique », ces données officielles numérisées sont, depuis le 1er décembre 2016, librement réutilisables en open data par tout un chacun : citoyen, startup, associations… »

Or, c'est faux. Les lois Valter et Lemaire imposent aux services d'archives d'accepter les demandes de réutilisation commerciale de leurs données, mais ceux-ci peuvent encore fixer des redevances.

Si les départements n'ont pas mis leurs licences en conformité avec la loi Valter au 1er décembre 2016, c'est en réalité à cause d'un téléscopage de lois. En effet, les services d'archives départementales attendaient alors un décret d'application relatif à la loi pour une République numérique. Cette loi prévoyait que les licences gratuites ne pourraient pas être mises en place par les collectivités de n'importe quelle manière, et qu'un décret expliquant le cheminement à poursuivre allait paraitre. Les collectivités attendaient donc la parution de ce décret avant de prendre une décision : opter ou non pour une redevance.

[1] Pierre-Alexandre Conte, l'affaire Filae : du besoin de repenser la politique de diffusion des données culturelles, décembre 2016 : http://www.lagazettedescommunes.com/478355/archives-laffaire-filae-du-besoin-de-repenser-la-politique-de-diffusion-des-donnees-culturelles/

Droits d'auteur réservés. OU (cc) BY-NC-ND

Certains acteurs, comme Lionel Maurel[1] appellent à « repenser la politique de diffusion des données culturelles[2] ». Celui-ci serait favorable à un partage à l'identique, qui n'empêche pas les réutilisations commerciales, mais qui impose aux acteurs privés la libre rediffusion des données[3].

Du côté de l'AAF, la question des redevances n'a pas à être soulevée. Son président rappelle que les licences payantes peuvent s'expliquer par le fait que « pour numériser, il faut payer quelqu'un qui numérise et qu'il y a des coûts liés à la conservation des documents, à leur mise à disposition et à leur mise à jour [4] ». Selon lui, le cas des archives est bien différent des données open datables, qui sortent d'applications et qui sont nativement numériques.

Cette affaire met en péril la démarche de numérisation des collectivités car l'administration pourrait reprocher aux services concernés de donner leur travail à des entreprises privées.

Cette affaire interroge sur la politique de diffusion des données des établissements culturels. Faut-il la repenser ? Pour Lionel Maurel, l'Open Data culturel se trouve aujourd'hui dans une « situation schizophrénique[5] ».

Tous les archivistes interrogés sont d'accord pour dire qu'il existe une « complexité juridique croissante[6] » et qu'une simplification serait la bienvenue.

Pour l'équipe VITAM, l'usager, comme les services publics d'archives, sont perdus face à cette complexité : « Plusieurs lois se recoupent sur l'accès en ligne : code du patrimoine, loi CNIL et loi CADA (et leurs pendants européens), droit d'auteur, de la propriété intellectuelle. Un document non accessible en ligne peut l'être en salle de lecture. Ce n'est pas toujours clairement connu d'un usager[7] ».

---

[1] Juriste et bibliothécaire, auteur d'un blog : https://scinfolex.com

[2] Lionel Morel : https://scinfolex.com/2016/12/13/une-enorme-faille-dans-la-loi-valter-sur-les-donnees-culturelles/

[3] Pierre-Alexandre Conte, l'affaire Filae : du besoin de repenser la politique de diffusion des données culturelles, décembre 2016 : http://www.lagazettedescommunes.com/478355/archives-laffaire-filae-du-besoin-de-repenser-la-politique-de-diffusion-des-donnees-culturelles/

[4] Ibid.

[5] Lionel Morel : https://scinfolex.com/2016/12/13/une-enorme-faille-dans-la-loi-valter-sur-les-donnees-culturelles/

[6] Voir annexe n°4.

[7] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

Dans ce deuxième chapitre, nous avons constaté que la règlementation sur l'accès aux archives publiques s'était adaptée au monde numérique. Nous avons vu également que le citoyen dispose de plusieurs outils pour étancher sa soif d'archives publiques. Mais comparativement à l'accès aux archives papiers, le numérique a-t-il engendré des transformations ?

Droits d'auteur réservés. OU (cc) BY-NC-ND

CHAPITRE III : UN ACCÈS TRANSFORMÉ

Ce chapitre est l'occasion de mettre en lien les deux chapitres précédents afin de constater les transformations qu'a provoqué l'accès aux archives publiques dans le monde numérique. Quels sont les grandes évolutions ? S'agit-il d'une véritable transition ou plutôt d'une adaptation à de nouveaux usages ?

Nous verrons que l'accès autour des données à caractère personnel correspond à une grande préoccupation, de la part du citoyen comme des institutions afin de garantir la protection de la vie privée. Comment concilier mémoire et droit à l'oubli (I) ? Puis nous remarquerons que l'accès via le numérique a grandement modifié les pratiques (II). Cet accès implique une nouvelle manière de rechercher l'information. Les mots clés vont-ils remplacer les instruments de recherche ? De plus, l'accès à un grand nombre de données publiques en temps réel sur les portails open data ont engendré une nouvelle pratique : la réutilisation. Enfin, nous illustrerons cette transformation du rapport à l'accès par le cas d'un service d'AD qui a placé l'innovation au centre de ses priorités afin de répondre à cette nouvelle demande d'accès de la part du public. Les services proposés par les AD doivent-ils évoluer ?

I> L'ACCÈS AUTOUR DES DONNÉES À CARACTÈRE PERSONNEL

Il convient tout d'abord de définir les données à caractère personnel. Celles-ci sont considérées comme étant à caractère personnel lorsqu'elles permettent d'identifier directement ou indirectement des personnes physiques. Par exemple, lorsqu'une personne est identifiée par son nom dans un fichier. Une donnée est également à caractère personnel si une personne est identifiable ou si un élément permet d'identifier indirectement une personne. Par exemple, un numéro de Sécurité sociale, de carte bancaire, une plaque minéralogique… Enfin, constituent également des données à caractère personnel toutes les données dont le recoupement permet d'identifier une personne précise.

La question qui nous est posée est la suivante : Internet est-il une mémoire virtuelle accessible sans limite ?

Nous traiterons tout d'abord de la réglementation européennes et nationale concernant les données à caractère personnel à travers notamment l'AU029 de la CNIL (A), puis nous

Droits d'auteur réservés. OU (cc) BY-NC-ND

nous intéresserons à la procédure et au droit de déréférencement (B) avant d'interroger des archivistes sur la tension entre mémoire et droit à l'oubli (C).

A> L'autorisation unique 029 et la règlementation européenne

Comme nous l'avons vu dans le chapitre précédent, les services publics d'archives se sont lancés dans la diffusion en ligne d'archives depuis les années 2000. À ce moment-là s'est posée la question des délais de mise en ligne. Quels délais respecter ? Plusieurs réponses ont été apportées par la communauté archivistique[1]. En effet, certains archivistes ont appliqué les délais de communicabilité déterminés par le code du patrimoine et que nous avons détaillés dans le premier chapitre de ce mémoire. D'autres n'ont mis en ligne que des documents de plus de 100 ou 120 ans.

C'est pourquoi la Commission nationale de l'informatique et des libertés a décidé d'encadrer et de réglementer les traitements des services d'archives relatifs à aux informations publiques contenant des données à caractère personnel[2]. Pour cela, la CNIL a deux volontés : d'une part, réglementer les délais de mise en ligne et mettre en place des processus d'occultation des données sensibles, et d'autre part, elle souhaite informer les personnes de leur droit d'opposition à cette diffusion en ligne.

L'autorisation unique 029, ou AU-029, a été élaborée par la CNIL avec le concours de la Direction générale des patrimoines du ministère de la Culture et rendue publique le 12 avril 2012. Celle-ci porte « autorisation unique de traitements de données à caractère personnel contenues dans des informations publiques aux fins de communication et de publication par les services d'archives publics[3] ».

L'AU 029 met en place plusieurs délais de mise en ligne[4], à ne pas confondre avec les délais de communicabilité car ils sont différents, de documents comportant des données à caractère personnel :

[1] Bruno Ricard, « L'AU 029 de la CNIL : cadre juridique de la mise en ligne des documents d'archives sur Internet », janvier 2015 : http://siafdroit.hypotheses.org/329

[2] https://www.cnil.fr/fr/declaration/au-029-archives-publiques

[3] Bruno Ricard, op.cit.

[4] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

- Il faut compter un délai de 25 ans pour la diffusion en ligne des actes de décès de l'état civil (les actes de décès sont immédiatement communicables dans un service public d'archives)
- 75 ans pour les actes de mariage de l'état civil (similaire aux délais de communicabilité)
- 75 ans pour les actes de naissance de l'état civil avec occultation des mentions marginales, ou 100 ans sans occultation de ces mentions (au lieu de 75 ans sans occultation des mentions marginales pour les délais de communicabilité)
- 100 ans pour les autres documents comportant des données à caractère personnel (au lieu de 50 ans ou de 25 ans après le décès pour les délais de communicabilité)
- 150 ans pour les « données sensibles » au sens de l'article 8 de la loi CNIL. Les données « sensibles » sont relatives aux origines raciales ou ethniques, aux opinions politiques, philosophiques ou religieuses, à l'appartenance syndicale, à la santé ou à la vie sexuelle des personnes, ou 100 ans si les documents ne sont accessibles qu'après création d'un compte-utilisateur. Ce compte permet de réguler l'accès et ne pas le permettre à tous. (Les délais de communicabilité concernent les données de santé sont de 120 ans après la naissance).
- L'AU 029 n'autorise pas la mise en ligne des documents comportant des données sensibles relatives aux infractions, condamnations et mesures de sûreté au sens de l'article 9 de la loi Informatique et Libertés. (Ces données sont communicables dans un délai compris entre 75 et 100 ans selon les cas dans un service public d'archives).

Enfin, la diffusion en ligne d'index nominatifs donnant accès aux documents numérisés n'est autorisée qu'après un délai de 120 ans. Tous les services d'archives sont tenus de respecter l'AU 029 qui possède une valeur prescriptive. Ils doivent notamment procéder à « engagement de conformité[1] » sur le site Internet de la CNIL. Tout comme pour les délais de communicabilité, des dérogations peuvent être délivrées par la CNIL au cas par cas. Un exemple de dérogation accordé par la CNIL est celui qui a permis la diffusion en ligne et l'indexation nominative des registres matricules relatifs aux soldats de la

[1] Bruno Ricard, op.cit.

Droits d'auteur réservés. OU (cc) BY-NC-ND

Première Guerre mondiale, dans le cadre du projet Grand Mémorial[1] (accordé en 2013 dans le cadre du centenaire de la Première Guerre mondiale).

Au-delà de l'AU 029, un règlement européen de mai 2016 vient rajouter une couche supplémentaire dans la protection des données à caractère personnel. Il s'agit du règlement 2016/679 relatif à la protection des personnes physiques à et à la libre circulation de ces données. Dans un contexte de lutte contre le profilage et de l'utilisation à outrance des données à caractère personnel par les GAFA, ce règlement a pour objectif de donner aux citoyens le contrôle des données qui les concernent. Plusieurs principes sont ainsi consacrés[2] :

- Le principe de loyauté. Les données sont traitées de manière licite et transparente.
- La limitation des finalités. Pour qu'une donnée soit collectée, il doit y avoir une finalité précise et légitime.
- La minimisation des données. Les données collectées doivent être limitées uniquement à celles nécessaires à la finalité.
- L'exactitude des données. Elles doivent être à jour et exactes.
- La limitation de la conservation. Les données ne doivent être conservées que pour le temps nécessaire à la finalité.
- L'intégrité et la confidentialité. Les données doivent être sécurisées.
- La responsabilité. Le responsable du traitement doit être capable de démontrer qu'il respecte bien toutes les obligations précédemment citées.

Tout individu concerné par des données à caractère personnel possède des droits. Il s'agit notamment du droit d'accès, du droit de rectification, du droit d'effacement (sous conditions). Des exceptions sont prévues, pour concilier protection des données à caractère personnel avec liberté d'expression.

L'exception majeure qui nous intéresse est celle concernant les « traitements à des fins archivistiques » pour les archives définitives. Grâce à cette exception, les données « sensibles » peuvent être conservées dans l'intérêt public, une dérogation au droit à l'oubli est reconnue, les données peuvent être conservées plus longtemps que la finalité

[1] Bruno Ricard, op.cit.

[2] Aude Roelly, « Données à caractère personnel : un nouveau règlement », mai 2016 : https://siafdroit.hypotheses.org/605#more-605

Droits d'auteur réservés. OU (cc) BY-NC-ND

initiale… Cela induit que pour les données à caractère personnel non sensibles ayant une finalité de valorisation du patrimoine, l'accès est autorisé sans restriction. Pour les données à caractère personnel sensibles ayant une finalité de valorisation du patrimoine à des fins historiques, scientifiques ou statistiques, l'accès doit être sécurisé. Par exemple, cet accès doit être justifié par un intérêt scientifique ou historique, et se faire via un compte utilisateur.

Pour Aude Roelly, cheffe du Bureau de la sélection, de la gestion et de la collecte au SIAF, « ces dérogations pour les traitements à des fins archivistiques étaient indispensables afin de ne pas mettre en cause la finalité même de la conservation des archives, à savoir la conservation de données fiables et intègres pour faire valoir des droits comme pour servir de matériaux à l'écriture de l'histoire[1] ».

B> Le déréférencement

En raison des risques croissants pour les droits fondamentaux des citoyens, et notamment pour le droit au respect de la vie privée, le traitement et la circulation des données à caractère personnel sur Internet sont une grande préoccupation actuelle.

Aujourd'hui, les grands moteurs de recherche d'internet de type Google sont tellement puissants qu'ils amassent un grand nombre de données sur chaque individu, sans que celui-ci en soit forcément conscient. Cet accès (parfois à outrance) à des données à caractère personnel peut parfois s'avérer contraire au respect de la vie privée. Le citoyen est donc en droit de se demander quelles informations le concernant sont en libre accès sur la toile ?

Il existe une manière de constater les informations qui circulent sur sa personne, et cela s'appelle « googliser ». En tapant son prénom et son nom dans un moteur de recherche, l'individu peut voir quels résultats apparaissent et quelles informations sont en libre accès concernant sa personne. Or, certaines informations peuvent lui être nuisibles. Par exemple, il n'est pas rare aujourd'hui qu'un employeur googlise des candidats à l'embauche afin de rechercher des informations sur ceux-ci.

Un nouveau droit pour le citoyen est apparu il y a seulement quelques années, afin de respecter sa vie privée : il s'agit du droit au déréférencement.

[1] Aude Roelly, op.cit.

Droits d'auteur réservés. OU (cc) BY-NC-ND

Voici une définition de la CNIL concernant ce droit[1] :

> Le déréférencement permet de demander à un moteur de recherche de supprimer certains résultats de recherche associés aux noms et prénoms. Il consiste à supprimer l'association d'un résultat de recherche à la requête "nom prénom". Cette suppression ne signifie pas l'effacement de l'information sur le site internet source. Le contenu original reste ainsi inchangé et est toujours accessible via les moteurs de recherche en utilisant d'autres mots clés de recherche ou en allant directement sur le site à l'origine de la diffusion.

Nous allons voir que c'est lors d'une affaire qui a opposé un citoyen espagnol à Google que cette règlementation européenne est apparue, consacrant notamment l'existence d'un droit à l'oubli numérique.

1) L'arrêt Google Spain

Ce droit est apparu en 2014. En se googlisant, un citoyen espagnol est tombé sur deux archives d'un quotidien espagnol relatant une saisie immobilière dans le cadre d'une procédure en recouvrement de cotisations sociales[2]. Afin de protéger sa vie privée, il a demandé à Google Spain (filiale espagnole) et Google Inc. à supprimer ou modifier ces informations qu'il jugeait dommageables pour sa personne. Il demande également à ce que ces résultats n'apparaissent plus dans la recherche Google.

Cette affaire est allée jusque devant la Cour de justice de l'UE qui a rendu son arrêt le 13 mai 2014. Cet arrêt se penche sur la directive 95/46/CE qui contient toutes les règles européennes relatives à la protection des données à caractère personnel. Cette directive est appliquée dans tous les États membres de l'UE.

Selon la CJUE, lorsqu'un internaute effectue une recherche en utilisant un moteur de recherche, « dès lors que l'exploitant de ce moteur de recherche extrait des informations, les enregistre et les organise, les conserve et, enfin, les communique aux internautes, il

[1] Site web de la CNIL : https://www.cnil.fr/fr/le-droit-au-dereferencement

[2] Article de presse : http://www.justice-en-ligne.be/article640.html

Droits d'auteur réservés. OU (cc) BY-NC-ND

réalise en réalité un « traitement » de données personnelles dont il est le « responsable » puisqu'il en détermine les finalités et les moyens »[1].

Donc, selon la CJUE, le moteur de recherche est bien attaquable en justice. La deuxième question qui se pose est de savoir si le moteur de recherche doit répondre à la demande de désindexation du citoyen espagnol. Ces informations, que le citoyen juge litigieuses, avaient fait l'objet d'une parution licite, le moteur de recherche avait seulement référencé ces contenus.

Mais la CJUE a bel et bien admis que ces informations devaient faire l'objet d'un déréférencement.

Pour elle, si la demande de désindexation répond à cinq critères[2], celle-ci doit être opérée par le moteur de recherche :

- Tout d'abord, l'information concernée doit faire l'objet d'une re-divulgation. Le moteur de recherche agit bien de la sorte en re-proposant l'information à la suite d'une recherche.
- Ensuite, la première divulgation de l'information doit avoir été effectuée licitement.
- Il faut également qu'un laps de temps important se soit écoulé entre les deux divulgations (16 ans pour le cas du citoyen espagnol, donc suffisant).
- L'information concernée ne doit pas faire l'objet d'un intérêt historique.
- Enfin, la CJUE déclare que ce « droit à l'oubli » ne peut concerner que des informations judiciaires. De plus, si l'information concerne un débat d'intérêt général ou une personne publique, ce droit ne peut pas être revendiqué.

La CJUE a donc admis la désindexation, mais non la suppression des informations. En réalité, la CJUE a cherché un compromis qui respecte deux droits fondamentaux du citoyen : d'une part, la nécessité de garantir un accès à l'information qui est un droit d'expression, et d'autre part, le droit à la vie privée.

[1] Article de presse : http://www.justice-en-ligne.be/article640.html

[2] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

Cette décision de la CJUE a donné lieu à de nombreuses critiques de la part de Google, mais aussi de Wikipédia qui parle de « censure du Web » ou encore de Reporters sans Frontières qui dénonce une décision « attentatoire à la liberté de l'information »[1].

2) Une législation encore à définir

Comme nous l'avons vu avec l'arrêt Google Spain, le droit au déréférencement n'est pas absolu. Il doit respecter un certain nombre de critère afin de garantir à la fois le respect de la vie privée et le droit d'accès à l'information. Cette nécessité de compromis fait qu'aujourd'hui le droit au déréférencement est très flou et nécessiterait une clarification.

En effet, en février 2016, agissant d'une manière similaire au citoyen espagnol, quatre individus ont demandé à Google le déréférencement de liens vers des informations les concernant[2]. Mais Google oppose à ces quatre personnes un refus. Ceux-ci saisissent alors la CNIL à la suite de ce refus. Mais la CNIL ne donne pas suite à leurs plaintes. Ils lancent alors un recours au Conseil d'État pour annuler ces décisions de clôture de leurs plaintes.

Les déréférencements souhaités par ces personnes concernait des liens vers des informations diverses : une vidéo montrant la relation d'une requérante avec une personne publique, un article de presse lié au suicide d'une adepte de l'Église de Scientologie, où il était fait mention qu'un requérant était alors chargé des relations publiques de cette organisation, des articles concernant une information judiciaire dans le cadre du financement d'un parti politique dont un requérant est impliqué, et enfin des articles concernant la condamnation d'un requérant pour agression sexuelle sur mineurs.

Pour le Conseil d'État, le droit de l'UE tel qu'interprété par la CJUE dans son arrêt Google Spain pose des difficultés d'interprétation. Le Conseil d'État se pose des questions quant au déréférencement d'informations sensibles, c'est-à-dire relatives à des opinions politiques, des orientations sexuelles, des opinions religieuses ou encore des condamnations judiciaires (comme tel est le cas pour ces quatre requérants). Quelles sont les obligations de déréférencement concernant des informations provenant d'articles de presse ou concernant des contenus inexacts ou incomplets ?

[1] Article de presse : http://convention-s.fr/decryptages/larret-google-spain-un-tournant-pour-le-droit-de-linformation/

[2] Site web du Conseil d'État : http://www.conseil-etat.fr/Actualites/Communiques/Droit-au-dereferencement

Droits d'auteur réservés. OU (cc) BY-NC-ND

Le Conseil d'État se déclare finalement comme n'étant pas en mesure de statuer, et demande une question préjudicielle. Ce procédé permet à une juridiction nationale d'interroger la CJUE sur l'interprétation ou la validité du droit de l'Union européenne dans le cadre d'un litige dont cette juridiction est saisie[1]. Une fois que la CJUE aura répondu, le Conseil d'État pourra alors rendre son verdict.

En définitive, cet exemple d'actualité nous montre à quel point le droit au déréférencement est un droit nouveau, qui n'a pas encore révélé tous ses aspects. La réponse de la CJUE va probablement modifier ce droit et peut-être imposer aux moteurs de recherche de nouvelles obligations ? Nous le saurons dans les mois à venir.

La tension entre mémoire et droit à l'oubli est bien ancrée, faut-il se souvenir plutôt qu'oublier ? Faut-il gommer plutôt qu'indexer ? Existe-t-il un compromis ?

C> La tension entre mémoire et droit à l'oubli

Nous avons souhaité interroger des archivistes sur cette question de la tension entre mémoire et droit à l'oubli[2]. Des réponses très variées nous sont parvenues, illustrant bien la tension qui existe au sein même de la profession.

Pour Cyril Longin, directeur des AM de Saint Etienne, les archivistes sont en dehors de cette question. L'archiviste est contraint par l'administration et par la loi : « pour la mise à disposition des ressources sur internet, la question ne se pose pas puisque qu'il y a la législation ». Il rajoute que les archivistes sont là pour permettre aux historiens et particuliers de conserver leur mémoire collective et individuelle tout en garantissant le droit individuel des personnes. Pour Cyril Longin, « il y a un droit à l'oubli, mais aussi un droit à l'histoire, et les deux ne sont pas contradictoires ». Pour lui il s'agit avant tout d'une question de gouvernance de l'information. Les archivistes sont les garants de la mémoire collective, « ils ont une certaine indépendance et distance par rapport à l'information qu'un particulier ou journaliste n'aura pas ». La sécurisation de l'accès est une solution pour concilier mémoire et droit à l'oubli selon lui.

Pour Frédérique Garnier, responsable du pôle des archives contemporaines aux AD de Loire-Atlantique, la réponse est beaucoup moins nuancée. Pour elle, « l'application du droit à l'oubli conduit à l'anonymisation des bases de données, instrument de recherche

[1] Site web du Conseil d'État : http://www.conseil-etat.fr/Actualites/Communiques/Droit-au-dereferencement

[2] Toutes les questionnaires sont disponibles en annexe de ce mémoire.

Droits d'auteur réservés. OU (cc) BY-NC-ND

par excellence des dossiers nominatifs, et outil indispensable de la recherche historique ». De son point de vue, le droit à l'oubli n'a donc pas sa place chez les archivistes.

Pour Marie Ranquet, conservateur du patrimoine et adjointe au chef du bureau du contrôle et de la collecte des archives publiques aux Archives de France, le régime actuel semble équilibré : Le principe de « ne pas diffuser en ligne de donnée à caractère personnel avant un long délai (délai de communicabilité + délai fixé par la CNIL : régime de l'AU29 de 2012, par exemple...) et de le doubler d'un accès plus court en salle de lecture » lui parait bien. Cependant selon elle, « certains délais en eux-mêmes pourraient être aménagés ».

Elle met tout de même en garde à ce que l'un des régimes ne prenne pas le pas sur l'autre, « car sinon, on ira forcément dans le sens d'une protection accrue, donc de délais plus longs pour tout le monde ». De son point de vue, mépriser la demande sociale de droit à l'oubli « risque d'être très contre-productif, car la réglementation en matière d'accès peut changer très vite, si elle est perçue comme pas assez protectrice ». Elle conclue en déclarant que « tenir compte du désir de protection est le seul moyen de continuer à être légitime dans notre politique d'accès aux archives : sinon on va nous reprocher d'être irresponsables... et remonter tous les délais, y compris en salle de lecture ».

Tout comme Marie Ranquet, Jeanne Mallet, conservateur du patrimoine et chargée de l'accès aux archives au SIAF, met aussi l'accent sur la demande sociale de l'accès aux archives : « cet accès simplifié et à grande échelle correspond à une vraie demande sociale : on voit notamment dans le cadre de projets comme le Grand mémorial un vrai désir d'avoir accès aux informations relatives aux soldats de la première guerre mondiale et d'écrire l'histoire de sa famille ».

De plus, elle met en exergue un paradoxe : « actuellement, nous voulons tout savoir sur tout le monde, mais nous souhaitons protéger au mieux nos propres données et pouvoir les détruire lorsque nous n'en avons plus besoin, ce qui est assez paradoxal quand nous sommes ravis de trouver les mêmes informations sur nos ancêtres ».

De son point de vue, la tension entre mémoire et droit à l'oubli « vient en réalité de l'absence de profondeur historique donnée à la réflexion : les règles qui valent pour les informations produites aujourd'hui cessent au bout d'un moment d'être pertinentes, notamment lorsque la personne concernée est décédée ». La mise en place de délais plus longs pour la consultation sur internet que pour la consultation en salle pour les documents comportant des données à caractère personnel, « garantit au mieux la

Droits d'auteur réservés. OU (cc) BY-NC-ND

protection des informations contenues dans les documents, qui ne sont mises en lignes qu'après le décès des personnes concernées ».

Enfin, pour Mélanie Rebours, conservateur du patrimoine et directrice de la diffusion et des partenariats du Programme VITAM, et pour son équipe, il est difficile d'apporter une réponse à cette question. « La donnée personnelle est un des produits phare des services d'archives. Les services d'archives permettent de garantir le droit à l'oubli dans le contexte de production tout en préservant, par l'archivage, la mémoire de l'information et sa réutilisation sur le long terme, dans le respect de la vie privée (entre autres) ». Pour ce membre de l'équipe, afin de se préserver du danger dû à la réutilisation des données à caractère personnel, une solution serait l'anonymisation : « le numérique offre des possibilités d'anonymisation non encore exploitées ». Selon cette personne, les services publics d'archives doivent avant tout sensibiliser le public : « Dans l'immédiat, je dirais que les services d'archives ont un rôle à jouer dans la sensibilisation aux problématiques d'accès aux données personnelles ».

Pour un membre de l'équipe VITAM, l'accès tel qu'il existe actuellement est satisfaisant. Pour cette personne ce sont les moteurs de recherche de type Google qui sont les principaux responsables : « on ne rend accessible en ligne que des archives à données personnelles avec des délais de communicabilité longs, voire on ne les rend accessible que via des plateformes sécurisées, non indexées par les moteurs de recherche externes. Ce sont ces derniers qui posent problèmes concernant le droit à l'oubli ».

Enfin, pour un dernier membre de l'équipe VITAM, « le droit à l'oubli est important en ce qui concerne des délais courts et des diffusions très larges d'informations sur Internet (il est plus critique en ce qui concerne la publication d'information sur les réseaux sociaux) ». Cette personne penche également vers l'anonymisation afin de rendre l'information disponible aux chercheurs : « Pourrait-on se permettre d'anonymiser des corpus pour les diffuser plus rapidement aux chercheurs sans prendre de risque ? C'est à mon sens une réflexion à mener afin de valoriser plus rapidement les fonds conservés ».

En définitive, nous pouvons retirer de cet échange plusieurs réponses à cette grande problématique qu'est la tension entre mémoire et droit à l'oubli : pour certains, la flexibilisation des délais de diffusion est à opérer afin de répondre à la demande sociale. Pour d'autres, l'anonymisation serait une solution. Dans tous les cas, ces archivistes conviennent qu'il faut sécuriser les données à caractère personnel via des accès particuliers.

Droits d'auteur réservés. OU (cc) BY-NC-ND

II> DE NOUVELLES PRATIQUES D'ACCÈS

Dans cette partie nous remarquerons que les archives numériques ont profondément transformé le rapport à l'accès. Est-ce l'archive qui doit s'adapter au web ou inversement ? Nous verrons que le numérique a engendré de nouvelles méthodologies de recherche archivistique liées à la pratique du web (A), de nouveaux usages liés aux données publiques disponibles en temps réel à travers la réutilisation (B) et enfin de nouveaux services chez les services publics d'archives dans le but de satisfaire la demande d'accès et d'autonomie du public virtuel (C).

A> L'impact du numérique sur la méthodologie de recherche archivistique

Si l'on se place du côté des internautes, effectuer une recherche via un moteur de recherche est devenu monnaie courante. Le moteur de recherche présente différents résultats sous la forme de liens, ce qui permet à l'internaute de naviguer aisément et de trouver rapidement l'information qu'il souhaite. Ce mode de recherche rend tous les sites égaux puisqu'ils apparaissent tous dans la liste de résultats. C'est aujourd'hui le mode de recherche privilégié des internautes. On voit mal comment on pourrait se passer des moteurs de recherche, quasiment toutes les recherches en ligne s'effectuent par un grand moteur de recherche.

Cette pratique est devenue chez les internautes une habitude de recherche. En effet, sans s'en rendre compte, nous avons progressivement acquis des codes. L'utilisation de moteurs de recherche en fait partie. C'est devenu un réflexe.

Or, la manière de rechercher des archives est radicalement différente et cela n'est pas forcément aisée à comprendre pour les néophytes. En effet, comme nous l'avons dit, les sites web proposent une navigation de liens en liens de niveaux équivalents. Un fonds d'archive, quant à lui, est structuré de manière arborescente, du niveau de description le plus large vers le niveau de description le plus précis[1]. Nous avons ici une logique verticale, et non horizontale comme proposée par le web.

De la même manière, le résultat recherché est différent d'une logique à l'autre. Avec une recherche web, l'internaute souhaite trouver directement du contenu, c'est-à-dire les

[1] Romain Wenz, « Quel rôle pour les portails dans le nouveau monde du web ? » (Article à l'état d'ébauche en vue d'une publication).

Droits d'auteur réservés. OU (cc) BY-NC-ND

informations qu'il cherche. A contrario, la logique verticale des archives conduit à trouver des références pour aller trouver l'information plus loin par la suite[1].

Une fois la recherche effectuée par le moteur de recherche, l'internaute s'attend à trouver l'information qu'il désire dans les premiers résultats, puisque ceux-ci sont classés par pertinence. Une fois de plus, la logique est différente pour les archives, puisque les instruments de recherche sont par nature neutres et impartiaux[2].

Avec les moteurs de recherche, l'internaute a pris l'habitude de taper dans un moteur de recherche des mots-clés associés au contenu final recherché. Par exemple, dans le secteur des bibliothèques, l'internaute va taper dans le catalogue directement le titre du livre, et non passer d'abord par les œuvres complètes. Dans le secteur des archives, les internautes qui souhaitent effectuer des recherches généalogiques vont commencer par taper dans le moteur le nom de leur aïeul, sans connaitre les sources d'informations (état civil, registre matricule…).

Or, si je recherche par exemple des informations sur la vie de Jules Verne à Nantes dans le moteur de recherche des AD44, et que je tape seulement Jules Verne, je vais aussi bien tomber sur des informations concernant le collège Jules Verne, la rue Jules Verne que sur des cartes postales représentant l'écrivain.

Cela fait partie des limites de la recherche par mots clés via des moteurs de recherche. Par exemple, contrairement à une recherche dans une salle de lecture, où le lecteur est accompagné et guidé par les archivistes, la recherche en ligne via un moteur de recherche comporte deux risques liés aux mots-clés utilisés par l'internaute. Si l'internaute utilise un mot-clé pas suffisamment précis, le moteur de recherche ne va pas lui fournir de réponses ciblées, efficaces. C'est ce qu'on appelle du « bruit ». Au milieu de cette masse d'informations, l'internaute est perdu et risque d'abandonner sa recherche, car trop laborieuse. Mais, au contraire, si le mot-clé choisi est trop précis, l'internaute a de grandes chances de voir apparaitre un message du type : « aucun résultat n'a été trouvé pour votre recherche ». La réaction de l'internaute sera de penser soit que l'information qu'il cherche n'est pas disponible en ligne, soit qu'elle n'existe pas. Il risque également d'abandonner sa recherche, alors qu'un autre mot-clé ou une recherche davantage ciblée aurait pu lui apporter un élément de réponse.

[1] Romain Wenz, op.cit

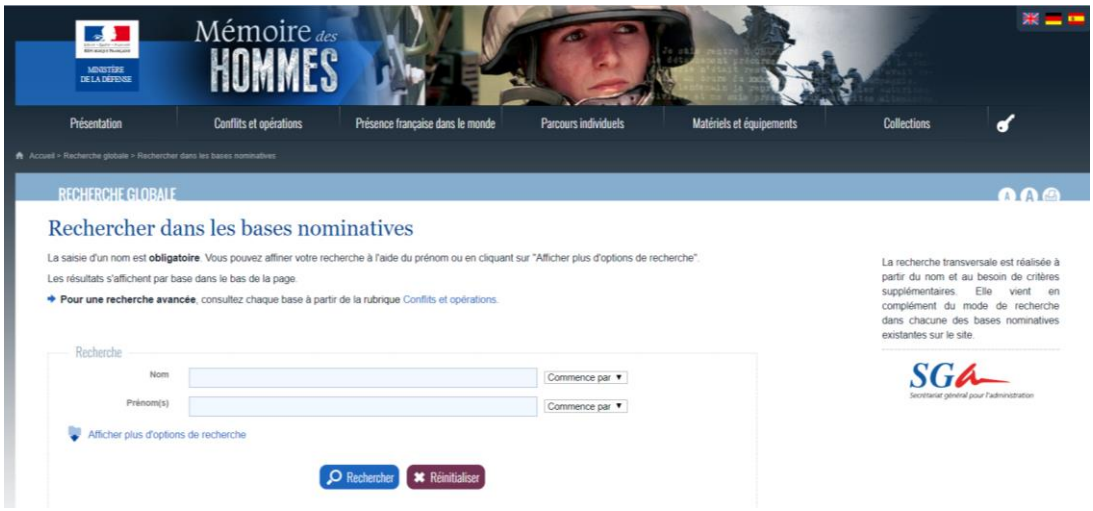
[2] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

Pour Romain Wenz, conservateur du patrimoine et chef de projet au SIAF, il s'agit d'une évolution structurelle des pratiques de recherche : « Nous sommes passés d'un modèle physique où il était naturel d'aller du général au particulier, des œuvres complètes à un livre précis, du registre au nom de personne, à un modèle virtuel où il est plus facile de commencer par des mots associés à la recherche finale, quitte à élargir ensuite la recherche[1] ».

Pour Mélanie Rebours et l'équipe Vitam, le mode de recherche archivistique doit évoluer : « Les travaux Vitam ont montré que les utilisateurs attendent un mode de recherche simple type « Google » : un champ permettant de saisir des mots clés[2] ».

C'est pourquoi les grands sites de généalogie ou des sites comme Mémoire des Hommes se sont aujourd'hui adaptés en proposant ce type de recherche précise.



Ci-dessus, le site Mémoire des Hommes, avec son moteur de recherche nominatif[3].

L'objectif de ces prochaines années pour les sites des services publics d'archives serait de réussir à relier les documents et les personnes, lieux ou thèmes qui peuvent être associés. L'idée est de donner à l'internaute des liens utiles, qui permettront de construire des services et une exploitation automatique dans des programmes. Romain Wenz nous donne un exemple simple et précis : « un internaute cherchant dans les archives le testament de Voltaire doit en fait trouver celui de François-Marie Arouet. Nous le savons,

[1] Romain Wenz, op.cit.

[2] Voir questionnaire

[3] Site web mémoire des hommes :

http://www.memoiredeshommes.sga.defense.gouv.fr/fr/arkotheque/client/mdh/recherche_transversale/bases_nominatives.php

Droits d'auteur réservés. OU (cc) BY-NC-ND

mais la machine doit apprendre à relier ces deux noms, ce que permettent aujourd'hui les outils du Web dit sémantique.[1] »

Le portail francearchives.fr va ainsi se développer dans cette lignée afin de rendre un service simplifié aux internautes. Les archivistes en charge de ce projet souhaitent que l'accès aux archives soit aussi intuitif que possible. Hervé Lemoine, directeur du SIAF, fait de l'accès une priorité : « dans une seconde étape, le portail continuera d'évoluer et de développer son offre de services innovants : cartes permettant de visualiser les lieux et documents de façon simple, frises chronologiques présentant les commémorations et événements en un clic, etc. Les outils du Web sémantique participeront au développement de ces services pour identifier en particulier les informations concernant les lieux, les dates et les thèmes ou événements[2] ».

Nous pouvons voir que cette méthodologie de recherche différente se retrouve également dans la construction même des instruments de recherche. Lorsqu'un site d'un service public d'archives met à la disposition de ses internautes des instruments de recherche sous forme numérique, ceux-ci sont en réalité bien différents des instruments de recherche papier. En effet, l'IR papier est davantage destiné à une lecture lente, posée, et donc celui-ci est généralement riche d'informations. On y trouve par exemple une introduction historique, une riche bibliographie, des sources complémentaires, des annexes, etc, tout cela répondant bien à la norme ISAD(G)[3]. Mais l'IR électronique quant à lui, est davantage destiné à une lecture rapide, afin de trouver l'information recherchée le plus vite. Toutes les informations précitées pour l'IR papier ne seront donc pas ou peu lues. A la place de ces informations, l'IR électronique est plutôt doté de liens pour que l'internaute puisse rebondir vers d'autres sources similaires. Pour l'équipe Vitam, il s'agit d'une vraie question : « faut-il faire des instruments de recherche ? La constitution de bases de données uniques est-elle réaliste, mais comment dans ce cas permettre la contextualisation qu'offrent les IR ? [4]»

En définitive, les sites web des services publics d'archive n'ont plus la seule vocation de pousser en ligne du contenu, il faut qu'ils arrivent à mettre en place un système pour que

[1] Romain Wenz, op.cit.

[2] Lettre du COEPIA (conseil d'orientation de l'édition publique et de l'information administrative, entretien avec Hervé Lemoine : http://www.gouvernement.fr/sites/default/files/contenu/piece-jointe/2016/12/lettre_du_coepia_ndeg_40.pdf

[3] Norme générale et internationale de description archivistique

[4] Voir questionnaire

Droits d'auteur réservés. OU (cc) BY-NC-ND

tout un chacun puisse trouver l'information qu'il recherche : la réponse adaptée semble être le web sémantique. En effet, pour Romain Wenz, « un portail ne peut plus aujourd'hui se limiter au rôle de « guichet d'entrée » pour la seule consultation de documents, fonction qu'avaient les portails des années 2000. Un portail fédérateur doit désormais assurer la diffusion maximale des données, et favoriser ce qu'on appelle la « réutilisation des données publiques », intégrée dans le mouvement Open data[1] ».

B> La réutilisation

En 2013, Aurélie Filippetti, alors ministre de la Culture et de la Communication, déclarait : « passer d'une politique de l'accès aux données à une politique de réutilisation des données : c'est le changement de paradigme qui est au cœur de la politique des usages numériques que je souhaite mener [2]».

La réutilisation des informations publiques est une utilisation par un tiers à d'autres fins que celles de la mission de service public pour laquelle les documents ont été produits ou reçus[3].

Ainsi, un exemple de réutilisation est celui de la société Filae qui exploite les documents d'état civil ou les registres paroissiaux à des fins commerciales.

Si nous faisons un focus sur le vocabulaire employé, la réutilisation porte sur les « informations publiques ». C'est-à-dire sur des informations publiées par des personnes publiques ou privées en charge d'une mission de service public. Mais tous les documents de ces organismes ne sont pas des informations publiques, il s'agit en réalité seulement des documents librement communicables, et sans droits de propriété intellectuelle.

Le CRPA pose le principe de la gratuité de la réutilisation. Il existe une dérogation à ce principe pour certaines administrations comme l'IGN ou Météo France, mais aussi pour les institutions – comme les services publics d'archives – qui lancent des opérations de numérisation[4]. Si un service public d'archives décide de tarifer la réutilisation, alors il doit choisir entre la licence ouverte et la licence ODbL.

[1] Romain Wenz, op.cit.

[2] Rapport sur l'ouverture et le partage des données publiques culturelles, décembre 2013 : http://cblog.culture.fr/rapport-ouverture-et-partage-des-donnees-publiques-culturelles/

[3] Bruno Ricard, « le nouveau régime juridique de la réutilisation des informations publiques », mai 2017 : https://siafdroit.hypotheses".org/tag/reutilisation

[4] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

Selon Simon Chignard, il existe plusieurs types de réutilisation[1]. Nous illustrerons chaque type de réutilisation par un exemple de réutilisation de données open data :

- La consultation : ce sont les données directement consultables par le grand public, le réutilisateur est l'utilisateur final. Il s'agit souvent de données froides et non complexes. Il s'agit simplement d'un jeu de données.

Par exemple, sur data.gouv.fr, la liste des monuments historiques[2]. Il s'agit bien de données froides et non complexes, que l'internaute va pouvoir réutiliser sans problème grâce notamment aux différents formats proposés (CSV, XML, JSON…).

- La médiation : les données brutes donnent lieu à des visualisations pour un public plus large, le réutilisateur est le médiateur de la donnée. Il s'agit souvent de données statistiques ou géographiques. La médiation permet une meilleure appréhension de la donnée. L'internaute dispose d'une interface qui lui permet de manipuler les informations. On parle également de data-journalisme pour ce type d'infographie.

Un exemple intéressant est celui de la cartographie des parrainages aux présidentielles 2017[3]. Cette réutilisation a été publiée le 15 mars 2017. Il s'agit d'une cartographie interactive des parrainages aux présidentielles 2017 réalisée à partir des données de

[1] CHIGNARD Simon, *Open data : comprendre l'ouverture des données publiques*, Limoges, Editions Fyp, 2012

[2] Portail data.gouv.fr : https://www.data.gouv.fr/fr/datasets/monuments-historiques-liste-des-immeubles-proteges-au-titre-des-monuments-historiques/

[3] http://serveur.arcorama.fr/parrainages2017

Droits d'auteur réservés. OU (cc) BY-NC-ND

Conseil Constitutionnel. L'internaute peut cliquer sur le candidat (en bas à gauche) et les données s'affichent immédiatement sur la carte.

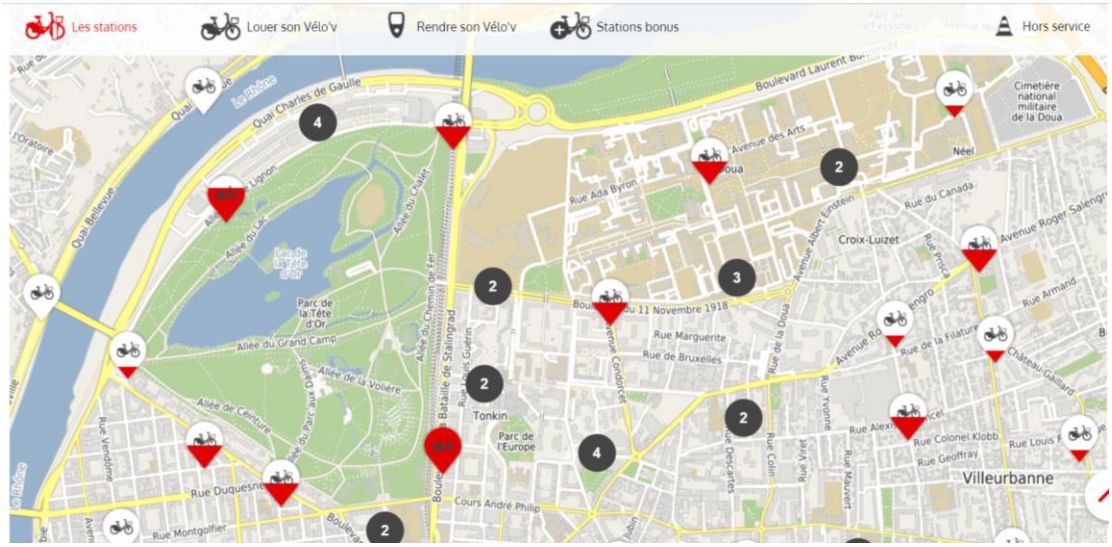
Ci-dessus, les parrainages de Marine le Pen en date du 22/03/2017 d'après les données officielles du Conseil Constitutionnel.

- Enfin, l'application : les données sont utilisées par les développeurs réutilisateurs pour construire des services et des applications pour les usagers finaux. Ici, ce sont des données chaudes, dynamiques. Elles possèdent des dimensions spatiales et temporelles précises.

Un bon exemple est celui de l'application Vélo'v[1] à Lyon qui permet de savoir en temps réel combien de vélos ou d'emplacements vélos en libre-service sont disponibles à telle ou telle station. Il s'agit bien de données chaudes et dynamiques car la carte présente la situation en temps réel. Grâce à son smartphone, un lyonnais peut ainsi gagner du temps et se diriger vers la station de vélos en libre-service la plus proche.

[1] https://velov.grandlyon.com/fr/les-stations.html

Droits d'auteur réservés. OU (cc) BY-NC-ND



Ci-dessus, les stations de vélos du parc de la tête d'Or à Lyon et de Villeurbanne. Une icône entièrement rouge correspond à une station qui est pleine. Au contrainte, une icône entièrement blanche montre une station totalement dépourvue de vélos.

En définitive, la réutilisation d'informations publiques donne à l'internaute une autre manière d'accéder aux données publiques. Selon la terminologie employée par Simon Chignard, la consultation, qui correspond à des données froides, n'apporte finalement pas d'informations supplémentaires à ce que pourraient apporter des archives papier (si ce n'est un accès facilité par internet). Mais on voit que la médiation apporte déjà un plus, un moyen de visualiser l'information. On note ici une transformation des chiffres bruts en informations immédiatement compréhensibles. Enfin, avec l'application, on n'assiste plus à une transformation de l'information mais bien à son exploitation à d'autres fins que celles exprimées initialement.

Les moyens d'accès aux données publiques sont ainsi décuplés et la donnée publique a permis de produire de la richesse par le développement d'applications. À travers l'étendue des possibles, on comprend mieux pourquoi la donnée est l'or noir du XXIème siècle.

C> Des services innovants : le cas des AD des Hautes-Alpes

Partant du constat que la numérisation des archives a eu pour effet d'éloigner le public des salles de lecture des services publics d'archives au profit de la consultation sur internet, Gaël Chenard, directeur des AD des Hautes-Alpes, a choisi de faire de la

Droits d'auteur réservés. OU (cc) BY-NC-ND

dématérialisation son leitmotiv. Selon lui, les archivistes doivent « repenser leurs pratiques de médiation pour accompagner le public virtuel[1] ».

1) Un public devenu virtuel

Entre 2005 et 2015, les AD les plus en pointe de la numérisation, ont perdu 50% de leurs lecteurs (59% des communications)[2]. L'étude des publics conduite par le SIAF en 2014 a révélé que ces internautes présentaient le même profil que celui des lecteurs physiques. Il s'agit donc d'une translation du public vers internet. Ce qui fait dire à Gaël Chenard : « lorsqu'ils ont le choix, nos lecteurs préfèrent ne pas avoir à se déplacer jusqu'à nous[3] ».

Afin de ne pas perdre tous leurs lecteurs physiques, les services d'AD tentent alors de mettre en avant leurs salles de lecture, comme des lieux de sociabilité. Mais pour le directeur des AD05, le pari est perdu d'avance : « Il est en réalité discutable qu'une salle de lecture, présumée silencieuse, soit un véritable lieu de sociabilité, contrairement à internet qui est aujourd'hui le réseau d'entraide et de partage le plus vaste du monde[4] ».

2) Un renversement du pouvoir

Cette transformation induite par le numérique provoque un autre phénomène. Avant que démarre véritablement cette révolution numérique, les archivistes étaient maitres et contrôlaient tout : conditions de consultation, horaires, accès aux originaux, priorités de numérisations... Aujourd'hui, ce rapport tend à s'inverser. Les internautes veulent décider des conditions dans lesquelles ils accèdent aux documents et les utilisent. Ils veulent également décider de ce qui doit être numérisé en fonction de leurs besoins ou de leurs envies. L'augmentation importante des demandes de recherches par correspondance (+35% entre 2005 et 2014[5]) est un des signes de cette évolution.

Ce qui se joue selon Gaël Chenard, « c'est l'émergence d'une consultation à distance qui devienne le mode normal de communication des documents. Les services d'archives ne sont pas isolés du reste des administrations, lesquelles avancent à grands pas vers l'e-

[1] Gaël Chenard, « L'acte II de la dématérialisation dans les archives », dans Culture et Recherche, n°134, les publics in situ et en ligne, hiver 2016-2017.

[2] Ibid.

[3] Ibid.

[4] Ibid.

[5] Ibid.

Droits d'auteur réservés. OU (cc) BY-NC-ND

administration. A bien y réfléchir, notre véritable public est déjà celui que nous ne voyons plus[1] ».

3) *Repenser le modèle d'accès et de recherche aux archives publiques*

Les AD05 ont construit leur site internet en 2014 avec comme objectif de mettre l'accent sur la dématérialisation complète de la communication. Ils ont ainsi souhaité anticiper la demande avant que celle-ci ne soit imposée par les lecteurs ou internautes. Leur objectif premier n'était pas de fermer leur salle de lecture physique, mais de proposer un service équivalent en ligne.

Le site internet des AD05 – www.archives05.fr – a été développé pour être le plus simple et épuré possible, en témoigne sa page d'accueil[2] :

Afin de proposer un service équivalent, il fallait répondre à deux enjeux majeurs que pose internet : accompagner le chercheur et lui permettre d'accéder à ce qu'il a besoin. Lui permettre d'accéder à ce dont il a besoin correspond à lancer des campagnes de numérisation. Tout d'abord, sont priorisés les fonds les plus consultés. Pour les archives consultées épisodiquement, la question de la rentabilité se pose. Économiquement, est-ce envisageable de numériser de tels fonds ?

Il serait plus rentable de numériser seulement selon le besoin plutôt que de financer des campagnes de numérisation systématiques. Les AD05 envisagent alors la numérisation à la demande comme une solution. Mais comment réussir à répondre la demande, si celle-

[1] Gaël Chenard, op.cit.

[2] https://www.archives05.fr/ consultée le 25/04/17

Droits d'auteur réservés. OU (cc) BY-NC-ND

ci est importante ? Pour les AD05, il est tout à fait possible d'anticiper la demande, par le biais d'études statistiques sur les documents consultés en salle de lecture. En effet, selon Gaël Chenard, « 75% des demandes des lecteurs se concentrent sur 8 à 10 % des fonds[1] ». Le projet des AD05 a donc consisté à numériser d'abord ces fonds prioritaires ou du moins leurs instruments de recherche de façon à pouvoir commander un document précis : 45% du fonds d'archives des AD05 est au moins référencé en ligne[2].

En décembre 2014, a été ouvert le service de numérisation à la demande. Son succès est sans conteste : 18 500 numérisations en 2015 (pour 95 000 numérisations à la demande sur l'ensemble des services départementaux cette année-là), et les chiffres seraient d'environ 40 000 pour 2016[3]. L'anticipation de la demande s'est révélée correcte puisque les demandes de numérisation portaient quasi exclusivement sur les archives demandées en salle de lecture.

Le coût de la numérisation reste abordable pour l'internaute : 0,25€ la page, ou un coût forfaitaire de 20€ pour les documents les plus volumineux. Les AD05 cherchent à inciter les internautes à faire reproduire la totalité du document. Gaël Chenard nous explique cette volonté : « nous partons du principe qu'il vaut mieux que l'utilisateur prenne tout le document et pas seulement quelques pages puisque de toute façon, il nous faut le manipuler et donc l'abîmer. Une fois qu'il est intégralement numérisé, c'est fait [4]». Le paiement peut s'effectuer directement en ligne. La numérisation à la demande aux AD05 correspond à trois ETP, ce qui montre l'importance dédiée à ce service.

Une fois cette problématique de l'accès résolue, les AD des Hautes Alpes ont souhaité tout mettre en œuvre afin de faciliter la recherche des internautes. Cette recherche n'est pas évidente comme le souligne le directeur des AD des Hautes-Alpes : « alors que beaucoup s'attendent à retrouver tout ce qui concerne leur grand père en tapant son nom dans un moteur de recherche, nos inventaires sont clairement rédigés pour retrouver l'information en suivant un parcours de recherche. Il faut donc concilier les deux en

[1] Gaël Chenard, op.cit.

[2] https://www.lemedia05.com/2015/27713/archives-departementales-unique-en-france-la-numerisation-a-la-demande-a-du-succes/

[3] Ibid.

[4] Gaël Chenard, op.cit.

Droits d'auteur réservés. OU (cc) BY-NC-ND

guidant les internautes le long de ce parcours, et ce de manière semi-autonome et aussi indolore que possible[1] ».

Pour cela, les AD05 ont décidé de construire sur leur site « une recherche pas à pas ». Ainsi, l'internaute n'a qu'à fournir les informations dont il dispose dans une interface logicielle, ce qui va permettre à celle-ci de lui présenter le parcours de recherche dont il a besoin. Le logiciel interroge les inventaires des archives en fonction de ces informations. Cependant, cette méthode ne fonctionne qu'avec des recherches récurrentes, comme l'état civil, les documents fonciers ou encore les registres matricules.

Pour des cas particuliers, les AD05 ont mis en place un service tout aussi innovant, à savoir la visioconférence avec un archiviste. L'internaute peut ainsi prendre rendez-vous directement sur le site, pour demander des conseils à un archiviste.

Une autre alternative est celle de regarder des tutoriels vidéo qui montrent comment effectuer certains types de recherche, qui demeure plus attractive et plus simple que de déchiffrer les fiches pratiques. Là encore, les AD05 se révèlent être les pionniers en France.

Les tutos sont disponibles sur Youtube via un lien sur le site des AD05. Ci-dessus, un tuto pour effectuer une recherche foncière[2].

L'objectif est de permettre à l'internaute de travailler en toute autonomie et d'approfondir sa pratique des fonds.

[1] Gaël Chenard, op.cit.

[2] Youtube : https://www.youtube.com/watch?v=o_fT6GBEQNk&list=PL1ZU_z_y41ZaEW46OVA3jPAzdzRk4spQ4

Droits d'auteur réservés. OU (cc) BY-NC-ND

Un autre objectif annoncé par les AD05 est d'impliquer l'internaute, le faire participer à l'activité des AD par des fonctions collaboratives. Cette démarche collaborative est très récente et tous les sites d'AD commencent à s'y intéresser. Pour le moment, les AD05 ne proposent aux internautes qu'un service d'annotations collaboratives. L'internaute est autorisé à laisser des commentaires sur les notices des instruments de recherche. Cela permet de partager les connaissances sur les fonds, d'interpeller la communauté d'utilisateurs qu'un document présente des renseignements inédits, ou encore qu'il parle d'un personnage important par exemple.

Bientôt, les AD05 proposeront à leurs internautes une rubrique d'indexation collaborative, comme c'est le cas déjà chez certains services publics d'archives. Par exemple, les AD44 proposent de participer à l'indexation des personnes répertoriées dans les recensements de population[1]. Un tutoriel vidéo explicatif décrit la manière de procéder.

Dans ce mouvement du « web collaboratif », on voit que la relation n'est plus unique comme elle était auparavant. Les AD n'ont plus la seule fonction de mettre à disposition des contenus pour les internautes. Ces sites cherchent à instaurer une relation à double sens, où leur public pourra s'impliquer, réagir, commenter, indexer ou participer. Le public des AD est ainsi encouragé à consommer du service à distance.

Enfin, concernant les demandes par correspondance, le parti pris des AD05 a été de répondre à toutes les demandes de recherche, afin notamment de mieux connaitre les demandes et de pouvoir une fois de plus les anticiper, afin de réduire à terme le nombre de ces demandes par correspondance.

Pour Gaël Chenard, la volonté recherchée est bien « un changement des modalités d'accomplissement de notre mission [d'archiviste] et non un simple empilement de services expérimentaux[2] ». Les internautes sont satisfaits, les réponses aux demandes sont effectuées dans des délais courts : 2 jours pour une numérisation, 4,5 pour une demande par correspondance. Mais il faut bien prendre en compte que le service des AD05 est un petit service, d'un petit département qui n'a par exemple aucun public universitaire. Mais

[1] Site web des AD44 : https://archives.loire-atlantique.fr/jcms/chercher/indexation-collaborative/indexation-collaborative-fr-t1_6181

[2] Gaël Chenard, op.cit.

Droits d'auteur réservés. OU (cc) BY-NC-ND

le site internet des AD, en proposant des services atypiques, modernes, avec une image jeune, montre bel et bien que les AD05 sont pleinement orientées numérique.

Gaël Chenard conclue de la manière suivante : « le site internet a été conçu comme un outil de reconquête autant que comme une réponse aux enjeux de toute la profession[1] ».

En définitive, nous constatons que c'est bien le monde des archives qui doit s'adapter au web et non l'inverse. Les IR, le travail de description archivistique, la méthodologie de recherche sont progressivement repensés pour s'adapter au numérique et répondre à la demande des internautes, qui réclament une plus grande autonomie. L'internaute s'est approprié le web, et il est gourmand en données et en archives. Pour le satisfaire, les outils d'exploration de contenus doivent s'affiner pour s'adapter notamment à la culture web. Enfin, il semble que la mise en relation des données via le web sémantique soit l'avenir des sites web des services publics d'archives.

[1] Gaël Chenard, op.cit.

Droits d'auteur réservés. OU (cc) BY-NC-ND

CONCLUSION

La problématique de ce mémoire était la suivante : comment se traduit l'évolution des formes d'accès aux archives publiques dans l'environnement numérique ?

Nous avons tout d'abord rappelé le contexte de la communication des archives, « au temps du papier ». Ce premier chapitre nous a donné les bases pour constater les évolutions apparues avec le numérique.

En effet, dans le deuxième chapitre nous avons constaté que le numérique a engendré de nouvelles formes d'accès : les portails et sites web se développent à vitesse grand V et les ressources en ligne ne cessent de s'accroitre. Un autre impact du numérique est la nécessaire adaptation de la règlementation. Cependant celle-ci nécessiterait une simplification. Nous avons également constaté que l'usager a pris possession du pouvoir. C'est lui désormais, de manière indirecte, qui décide de ce qui doit être diffuser en ligne en priorité. L'État et les collectivités ont bien compris l'enjeu que représente la diffusion des données en ligne, et ceux-ci jouent le jeu de la transparence. Pour les services publics d'archives, cet accès en ligne est devenu un objectif stratégique prioritaire. De plus, ils ont perdu leur monopole puisque de nouveaux acteurs sont apparus tel que des sociétés privées ou des administrations et établissements publics.

Enfin, dans un dernier chapitre, nous nous sommes intéressés à l'accès aux données à caractère personnel, qui font débat dans notre société. Une nouvelle législation est en cours d'élaboration afin de garantir la vie privée des citoyens, et de permettre à la fois le droit à l'information dont ils disposent. Cependant, la tension entre mémoire et droit à l'oubli est toujours présente. Enfin, en nous penchant sur le cas des AD05, nous avons constaté que les services publics d'archives proposent de tout nouveaux services aux internautes. C'est le métier même de l'archiviste qui évolue avec le numérique.

Nous avons souhaité connaitre le point de vue des archivistes quant aux perspectives d'évolution de l'accès aux archives publiques. La question posée était

Droits d'auteur réservés. OU (cc) BY-NC-ND

la suivante : comment voyez-vous l'accès de demain ? Ceux-ci nous ont répondu de manières diverses et variées, offrant une ouverture intéressante à ce mémoire[231].

Cyril Longin entrevoit plusieurs perspectives d'évolution. Pour le directeur des Archives municipales de Saint-Etienne, « la question des sites d'archives est à mettre en lien avec les sites d'Open Data. Nos sites internet seront assez proche des plateformes open data. Nous allons devenir le réceptacle de données produites, un véritable entrepôt de données ». Enfin selon lui, l'approche du passage de la communication à la diffusion irait encore plus loin, car les services publics d'archives auront « obligation de diffuser ».

Frédérique Garnier, responsable du pôle des archives contemporaines aux AD de Loire-Atlantique nous donne de nombreuses pistes comme par exemple le web participatif : « identification de documents, notamment photographiques. Enrichissement des descriptions : bases de données nominatives », le web sémantique, la mutualisation des catalogues, la réutilisation « via des outils de communication de façon à créer des *timeline* ou *storymap* », et enfin « l'utilisation du web social pour offrir une plus grande visibilité ».

Pour Marie Ranquet, il faut « continuer à augmenter et accompagner les possibilités techniques d'accès à distance (accès à distance à des archives non légalement diffusables, par exemple, par le biais d'espaces réservés sur les sites internet, en sorte d'intranet distant pour lecteurs ? possibilité notamment pour les archives consultables par dérogation ?) » Mais elle rajoute qu'il ne faut pas être naïf : « ce n'est pas parce que le ministère de l'Intérieur (par exemple !) joue la vertu en publiant des données en open data que certaines de ses archives sensibles seront plus facilement accessibles... voir la difficulté qu'on a eue à obtenir l'ouverture des archives de la Seconde Guerre mondiale ! »

Enfin selon elle, il y aurait un chantier à mener sur la simplification du droit d'accès : « Les régimes spéciaux d'accès aux archives publiques pour lesquelles une diffusion particulière a été jugée nécessaire se sont multipliés (cas de l'accès aux archives électorales, sociales, judiciaires, environnementales…). L'on observe aussi une multiplication de régimes d'accès anticipé avant expiration des délais de communicabilité, tendance due notamment aux inflexions nouvelles qui tendent à

[231] Voir en annexe

Droits d'auteur réservés. OU (cc) BY-NC-ND

favoriser l'ouverture des données publiques. Le nombre de ces régimes rend l'ensemble de l'édifice du droit d'accès à l'information publique très difficilement compréhensible pour les administrations et a fortiori pour les administrés, qui ne peuvent que pâtir d'une telle situation. Il faudrait ainsi recenser tous les régimes d'accès et envisager leur refonte globale ».

Pour Jeanne Mallet, les délais de mise en ligne sont appelés à évoluer dans le sens d'un raccourcissement : « Les usagers souhaitent en réalité la diffusion des documents qui concernent leur famille afin de pouvoir faire leurs recherches plus aisément. Seule une demande de retrait a eu lieu depuis le lancement du Grand mémorial : il y a donc une acceptation sociale de la diffusion des informations relatives aux grands-parents et au-delà, lorsqu'ils sont décédés ».

Elle souhaite également que l'accès soit encore simplifié, « en permettant un accès distant généralisé. Je ne pense pas que cela entraîne une disparition des salles de lectures, et il ne s'agirait même dans ce cas que d'une rematérialisation d'une salle de lecture distante, et non d'une disparition. Enfin selon elle, « les projets permis par le numérique, notamment en termes de valorisation, sont également appelés à se développer ».

Pour Mélanie Rebours et l'équipe Vitam, les perspectives d'évolution sont « le développement des téléservices en ligne : hotline, chat, conseil à distance, une meilleure formation des internautes sur les outils disponibles en ligne et en interne (au sein des services d'archives), sur les différents droits couvrant les données (personnelles) mises en ligne, une simplification des droits sur l'accès et une nouvelle la méthode de recherche : utiliser un moteur de type « google » ? »

L'accès aux archives publiques a donc considérablement évolué avec le numérique. La pratique archivistique est en phase d'adaptation, et les archivistes ont bien le pied à l'étrier du numérique pour répondre aux nouveaux besoins.

Droits d'auteur réservés. OU (cc) BY-NC-ND

SOURCES

Mes sources sont constituées par l'enquête que j'ai réalisée auprès de différents archivistes.

Voici les archivistes qui ont répondu à mon enquête :

- Cyril Longin, directeur des AM de Saint-Etienne.
- Frédérique Garnier, conservateur du patrimoine et responsable du pôle des archives contemporaines aux AD de Loire-Atlantique.
- Marie Ranquet, conservateur du patrimoine et adjointe au chef du bureau du contrôle et de la collecte des archives publiques aux Archives de France.
- Jeanne Mallet, conservateur du patrimoine et chargée de l'accès aux archives au SIAF.
- L'équipe VITAM, représentée par Mélanie Rebours, conservateur du patrimoine et directrice de la diffusion et des partenariats du Programme VITAM.

Ci-dessous les 6 questions posées lors de cette enquête :

1) Quelle est votre définition de l'accès aux archives publiques ?
2) En quoi le numérique est-il venu modifier cet accès ?
3) Quels sont les avantages et inconvénients de l'accès via le numérique ?
4) Quels moyens votre service met-il en œuvre concernant l'accès numérique ?
5) Quel accès dans la tension entre mémoire et droit à l'oubli ?
6) Comment voyez-vous l'accès de demain (perspectives d'évolution) ?

Les questionnaires complétés sont disponibles en annexe.

Droits d'auteur réservés. OU (cc) BY-NC-ND

BIBLIOGRAPHIE

Ouvrages de référence :

- BANAT-BERGER Françoise, DUPLOUY Laurent, HUC Claude, *L'archivage numérique à long terme : les débuts de la maturité ?*, Paris, la Documentation française : direction des archives de France, 2009, 284 p.
- CHAVE Isabelle (sous la direction de) et collectif d'auteurs, *Abrégé d'archivistique, principes et pratiques du métier d'archiviste*, Paris, AAF, 2012, 346 p.
- COEURE Sophie, DUCLERT Vincent, *Les archives*, Paris, La découverte, 2011, 126p.
- FAVIER Jean, *les archives*, PUF, que sais-je ?, 2001, 127p.
- GALLAND Bruno, *Les archives*, Paris, PUF, 2016, 127 p.

Thématique de l'accès aux archives numériques :

- ROUSTAN Mélanie (sous la direction de) et collectif d'auteurs, *La recherche dans les institutions patrimoniales : sources matérielles et ressources numériques*, Villeurbanne, Presses de l'enssib, 2016, 198 p.

Open data :

- BOURCIER Danièle, FILIPPI Primavera de, *Open data & Big data : nouveaux défis pour la vie privée*, Paris, Mare & Martin, 2016, 269 p.
- CHIGNARD Simon, *Open data : comprendre l'ouverture des données publiques*, Limoges, Editions Fyp, 2012, 191 p.
- FAVRO Karine, *Open data : une révolution en marche*, Paris, Victoires éditions, 2016, 127 p.
- GOUIGOUX Jean-Philippe, *Open data : consommation, traitement, analyse et visualisation de la donnée publique*, St Herblain, Editions ENI, 2016, 583 p.
- GUERRE Louise, *Open data en France : acteurs, projets et tendances*, Paris, Serdalab, 2012, 226 p.
- KITCHIN Rob, *The data revolution : big data, open data, data infrastructures and their consequences*, Los Angeles, 2014, 222 p.

Droits d'auteur réservés. OU (cc) BY-NC-ND

- LEJEUNE Yannick et collectif d'auteurs, *Big, fast, open data : décrire, décrypter et prédire le monde : l'avènement des données*, Editions Fyp, 2014, 191 p.
- MONINO Jean-Louis, *Big data, Open data et valorisation des données*, Londres, ISTE Editions, 2016, 158 p.
- PAQUIENSEGUY Françoise, *Open data : accès, territoires, citoyenneté : des problématiques info-communicationnelles*, Paris, Editions des archives contemporaines, 2016, 141 p.

Mémoires :

- AGOSTINI Clémence, *L'ABES et la BnF en route vers l'open data*, sous la direction de Gildas Illien, ENSSIB, 2015.
- GAILLARD Rémi, *De l'Open data à l'Open research data : quelles politiques pour les données de recherche ?*, sous la direction de Monique Joly, ENSSIB, 2013.
- VENTRESQUE Vincent, *Les mutations des collections numériques à l'heure du web de données*, sous la direction de Françoise Banat-Berger, ENSSIB, 2013.

Thèse :

- RANQUET Marie, *L'accès aux archives publiques en France. Le droit et la pratique vus par les archivistes depuis 1979*, sous la direction de Christine Nougaret, ENC, 2016.

Revues :

- Document numérique, *Accès à l'information et évaluation*, Lavoisier, 2015/1 (Vol. 18), 124 p.
- Document numérique, *Variabilité et adaptation de l'accès à l'information*, Lavoisier, 2014/2 (Vol. 17), 148 p.

Sitographie :

- Blog du SIAF *Droit(s) des archives* : http://siafdroit.hypotheses.org/
- Légifrance : https://www.legifrance.gouv.fr/
- Portail France Archives : https://francearchives.fr/fr/

Droits d'auteur réservés. OU (cc) BY-NC-ND

- Site de la CADA : http://www.cada.fr/
- Site de la CNIL : https://www.cnil.fr/

Droits d'auteur réservés. OU (cc) BY-NC-ND

ANNEXES

Table des annexes

ANNEXE 1 : ENQUÊTE AUPRÈS DE FRÉDÉRIQUE GARNIER 122
ANNEXE 2 : ENQUÊTE AUPRÈS DE MARIE RANQUET 124
ANNEXE 3 : ENQUÊTE AUPRÈS DE JEANNE MALLET.....127
ANNEXE 4 : ENQUÊTE AUPRÈS DE L'ÉQUIPE VITAM.....131
ANNEXE 5 : ENQUÊTE AUPRÈS DE CYRIL LONGIN.....137

Droits d'auteur réservés. OU (cc) BY-NC-ND

ANNEXE 1 : ENQUÊTE AUPRÈS DE FRÉDÉRIQUE GARNIER

1) Quelle est votre définition de l'accès aux archives publiques ?

La communication des archives (et donc l'accès aux archives) fait partie des missions des services d'archives. Collecter et conserver ne sert à rien si les documents ne sont pas communiqués. L'accès aux archives est un droit, maintenant indissociable de l'accès au document administratif (L300-1 et suivant du code des relations entre le public et l'administration) et des délais de communicabilité des archives (art L213 et suivant du code du Patrimoine).

2) En quoi le numérique est-il venu modifier cet accès ?

La numérisation des archives, la constitution d'images numériques initialement envisagées comme le support de substitution succédant aux microfilms, est devenu, avec la possibilité technique de diffuser massivement ces images en ligne, un formidable outil de communication. Les salles de lecture virtuelles touchent des dizaines de milliers d'internautes là où les salles de lecture physique n'en touchent que quelques milliers. Par ailleurs, l'accès est immédiat et peut toucher des publics empêchés. Il peut également se faire sans contrainte horaire et ne nécessite pas de déplacement.

3) Quels sont les avantages et inconvénients de l'accès via le numérique ?

Les avantages sont essentiellement liés à la diffusion (cf. plus haut). On peut également envisager une appropriation plus grande des archives.

Parmi les inconvénients, on pourrait mettre en avant le fait que, si la numérisation met en lumière quelques fonds, souvent déjà les plus consultés, toutes les archives ne peuvent être numérisées et la plus grande partie des documents restent et resteront dans l'ombre. Par ailleurs, on ne peut que constater l'impact de la mise en ligne des fonds sur la fréquentation des salles de lecture physique (en baisse, même si ce n'est pas le seul facteur d'explication). L'aspect médiation disparait : l'archiviste ne peut plus aider à la recherche le lecteur seul derrière son écran. On peut craindre également une rupture d'égalité de traitement / une fracture numérique (haut débit, manque de connaissances informatiques, niveau d'équipement). Enfin l'habitude d'accéder à des documents numérisés engendre parfois une frustration chez les

Droits d'auteur réservés. OU (cc) BY-NC-ND

usagers lorsque les archives ne le sont pas ou que la consultation des inventaires ne débouche pas sur une communication immédiate des documents.

La communication est entièrement tributaire de la technique : si le site internet est hors service, le service est interrompu. Enfin, on assiste depuis quelques années à un phénomène de marchandisation des images d'archives et des bases de données permettant d'y accéder (notamment par des sociétés faisant commerce de la généalogie récréative). L'accès numérique pose également des contraintes techniques pour le service d'Archives et les usagers (pérennité des formats, versions de navigateurs, anti-virus, outils de visualisation, stockage, environnement du poste de travail…)

4) Quels moyens votre service met-il en œuvre concernant l'accès numérique ?

Site internet, numérisation des fonds (30 000 euros par an pour la numérisation, 20 000 euros pour la mise en ligne et le stockage des images), inventaires des fonds

5) Quel accès dans la tension entre mémoire et droit à l'oubli ?

L'application du droit à l'oubli conduit à l'anonymisation des bases de données, instrument de recherche par excellence des dossiers nominatifs, et outil indispensable de la recherche historique.

6) Comment voyez-vous l'accès de demain (perspectives d'évolution) ?

Les sirènes du web 2.0 parfois difficiles à mettre en œuvre c'est-à-dire participation des internautes à l'enrichissement de l'offre (identification de documents, notamment photographique. Enrichissement des descriptions : base de données nominative, transcription… ?).

L'utilisation du web sémantique et des métadonnées pour la recherche dans les archives nativement numériques.

Protocoles OAI PMH, catalogues mutualisés, développer la réutilisation des données via des outils de communication(ex : timeline, storymap…) utiliser le web social afin d'avoir le maximum de visibilité (facebook, twitter….)

Droits d'auteur réservés. OU (cc) BY-NC-ND

ANNEXE 2 : ENQUÊTE AUPRÈS DE MARIE RANQUET

1) Quelle est votre définition de l'accès aux archives publiques ?

Avant tout une question de démocratie et d'exercice des droits. Le côté "ludique" ou "recherche", même scientifique, passe à mon avis au second plan. Le grand défi est là, dans la mesure où ce sont ces aspects qui sont le plus souvent mis en avant par les politiques culturelles du fait de leur succès : ne pas oublier qu'avant les chercheurs, avant les généalogistes, il y a des personnes qui ont un besoin vital d'accéder aux archives. Or la législation donne l'impression de s'attacher davantage à l'aspect recherche qu'à l'aspect exercice des droits (voir la loi pour une République numérique par exemple) : du coup on fait de l'open data à tout va, mais ce faisant, on néglige l'accès traditionnel, sur pièces. Or la plupart des archives ne sont pas "opendatables"... et l'accès à celles-ci a plutôt tendance à se durcir. (voir l'article que j'ai écrit avec Aude Roelly : « Faut-il euthanasier les archives ? : tension entre mémoire et oubli dans la société française contemporaine », La Gazette des archives, t. 2017/1 Meta/morphoses. Les archives, bouillons de culture numérique, n° 245, 2017, p. 139-154.)

2) En quoi le numérique est-il venu modifier cet accès ?

3) Quels sont les avantages et inconvénients de l'accès via le numérique ?

J'y vois un malentendu. On est partis d'une formidable opportunité, dématérialiser l'accès, permettre l'accès distant aux archives, etc., qui sont vraiment des grandes avancées, indéniablement. Mais en cours de route, cette "facilité" donne l'impression de prendre le pas sur l'accès "brut" aux archives sans filtre, sans anonymisation, sans délai supplémentaire que le simple délai de communicabilité... or il est primordial, pour la sauvegarde des droits des personnes ainsi que pour l'exercice démocratique, de pouvoir continuer à assurer un accès sans filtre aux archives. J'entends par filtre les conditions réglementaires qu'on impose pour la diffusion en ligne des archives : pas de donnée à caractère personnel, délai supplémentaire pour la diffusion. Il faut que ces deux modalités d'accès coexistent, sans que l'une prenne le pas sur l'autre... mais c'est pourtant ce qui est en train de se passer, à cause de (grâce à ?) l'effet de mode "open data". Qui est en lui-même une bonne chose, bien sûr...

Droits d'auteur réservés. OU (cc) BY-NC-ND

4) Quels moyens votre service met-il en œuvre concernant l'accès numérique ?

Mise en place du portail national Francearchives.fr pour fédérer l'ensemble des ressources archivistiques disponibles à la fois en ligne et en salle de lecture (renvoi vers documents numérisés + instruments de recherche).

5) Quel accès mettre en place dans la tension entre mémoire et droit à l'oubli ?

Le régime actuel me semble équilibré : ne pas diffuser en ligne de donnée à caractère personnel avant un long délai (délai de communicabilité + délai fixé par la CNIL : régime de l'AU29 de 2012, par exemple... même si certains délais en eux-mêmes pourraient être aménagés. C'est le principe qui me paraît bien.), mais le doubler d'un accès plus court en salle de lecture. Faire très attention que l'un des régimes ne prenne pas le pas sur l'autre, car sinon, on ira - forcément - dans le sens d'une protection accrue, donc de délais plus longs pour tout le monde. Il faut absolument faire attention à la demande sociale de "droit à l'oubli": la mépriser (et donc, mettre en ligne trop vite, par exemple) risque d'être très contre-productif, car la réglementation en matière d'accès peut changer très très vite, si elle est perçue comme pas assez protectrice. Tenir compte du désir de protection est le seul moyen de continuer à être légitime dans notre politique d'accès aux archives : sinon on va nous reprocher d'être irresponsables... et remonter tous les délais, y compris en salle de lecture.

6) Comment voyez-vous l'accès de demain (perspectives d'évolution) ?

Point à relier au précédent : se méfier d'un durcissement de la réglementation, déjà vérifié ; en même temps, continuer à augmenter et accompagner les possibilités techniques d'accès à distance (accès à distance à des archives non légalement diffusables, par exemple, par le biais d'espaces réservés sur les sites internet, en sorte d'intranet distant pour lecteurs ? possibilité notamment pour les archives consultables par dérogation ?) Mais ne pas être naïf : ce n'est pas parce que le ministère de l'Intérieur (par exemple !) joue la vertu en publiant des données en open data que certaines de ses archives sensibles seront plus facilement accessibles... voir la difficulté qu'on a eue à obtenir l'ouverture des archives de la Seconde Guerre mondiale !

Par ailleurs il y a un chantier à mener sur la simplification du droit d'accès. Les régimes spéciaux d'accès aux archives publiques pour lesquelles une diffusion

Droits d'auteur réservés. OU (cc) BY-NC-ND

particulière a été jugée nécessaire se sont multipliés (cas de l'accès aux archives électorales, sociales, judiciaires, environnementales…). L'on observe aussi une multiplication de régimes d'accès anticipé avant expiration des délais de communicabilité, tendance due notamment aux inflexions nouvelles qui tendent à favoriser l'ouverture des données publiques. Le nombre de ces régimes rend l'ensemble de l'édifice du droit d'accès à l'information publique très difficilement compréhensible pour les administrations et a fortiori pour les administrés, qui ne peuvent que pâtir d'une telle situation. Les administrations, souvent dans l'ignorance des deux grands textes précités, n'ont pas cessé ces dernières années de tenter de prendre des dispositions particulières s'agissant des informations publiques nativement numériques alors même que des dispositifs réglementaires préexistent. Ce foisonnement du droit et la méconnaissance du rôle du ministère de la culture ont rendu possible la création de commissions ad hoc pour tenter de résoudre les difficultés nées de la multiplication de textes de niveaux très différents. Il faudrait ainsi recenser tous les régimes d'accès et envisager leur refonte globale...

Droits d'auteur réservés. OU (cc) BY-NC-ND

ANNEXE 3 : ENQUÊTE AUPRÈS DE JEANNE MALLET

1) Quelle est votre définition de l'accès aux archives publiques ?

L'accès est la justification de la collecte et de la conservation des documents : il n'est utile de conserver que dans la mesure où des citoyens ont ensuite la possibilité de consulter. Il s'agit donc de permettre à toute personne qui en justifie le besoin, ou simplement l'envie, la consultation de documents utiles à son histoire personnelle, à ses recherches scientifiques ou qui lui sont nécessaires pour prouver un droit. Ce dernier aspect est d'ailleurs, historiquement et juridiquement, la justification première du droit d'accès aux archives tel qu'il est conçu en France. C'est un rouage essentiel de notre démocratie.

2) En quoi le numérique est-il venu modifier cet accès ?

Le numérique a permis un accès plus général et donc une meilleure connaissance du droit des citoyens à avoir accès aux documents. Cette évolution est liée fortement à l'engouement pour la généalogie et à la numérisation par les services d'archives de milliards d'images, principalement d'état civil : les recherches qui nécessitaient autrefois des déplacements sont aujourd'hui très nettement simplifiées. Cela a permis notamment une très grande visibilité des services d'archives, qui sont pour le niveau départemental les sites institutionnels les plus consultés d'assez loin (je n'ai plus les chiffres exacts en tête, mais il s'agit de millions de visiteurs uniques par an). Le numérique permet donc une réelle mise en œuvre du principe selon lequel tout citoyen peut avoir accès aux informations le concernant ou concernant l'État, qui auparavant était plutôt théorique car perçu comme très complexe. Je précise cependant que c'est surtout l'échelle qui a été modifiée par le numérique : les principes sont toujours les mêmes et n'ont pas subi de profonde modification, en ce qui concerne l'accès, depuis la Révolution.

Les évolutions du droit ont également pris en compte cette possibilité d'accès, en inscrivant dans la loi la possibilité pour les usagers de choisir les modalités d'accès aux documents (article L. 311-9 du code des relations entre le public et l'administration), qui peut être dématérialisé au choix du demandeur. Le numérique a enfin, à mon avis, un impact sur la façon dont les usagers perçoivent l'accès. Dans la mesure où il est aujourd'hui possible de transmettre "instantanément" et "sans frais" les documents à l'autre bout de la planète, les

Droits d'auteur réservés. OU (cc) BY-NC-ND

personnes qui souhaitent avoir accès aux documents comprennent mal les obligations qui leur sont parfois faite de se déplacer, voire de payer les reproductions des documents qu'ils demandent lorsqu'ils sont numérisés pour permettre leur envoi. Le numérique peut donc également avoir un aspect négatif, lié surtout à une méconnaissance des enjeux notamment financiers qu'il représente pour les services.

3) Quels sont les avantages et inconvénients de l'accès via le numérique ?

J'ai partiellement répondu à cette question dans ma réponse précédente. Les avantages sont la simplification pour les lecteurs et la visibilité accrue des services, comme déjà évoqué.

L'inconvénient est une mauvaise compréhension de la façon dont le numérique fonctionne, notamment la croyance paradoxale en la gratuité du numérique (alors que les campagnes de numérisation se calculent en millions d'euros) de la part du public. Cette croyance se trouvent également chez les décideurs, puisque les débats autour de la loi pour une république numérique à l'Assemblée nationale ont notamment comporté des échanges sur le sujet : une fois que les documents sont numérisés, ou lorsqu'ils sont nativement numériques, les décideurs ne prennent pas forcément en compte les coûts de conservation et de communication qui sont pourtant bien réels.

Enfin, le dernier inconvénient du numérique est justement sa très grande facilité d'usage et d'accès : la conservation des documents nativement numériques pose nombre de questions, et la facilité avec laquelle ils sont détruits pose des difficultés en terme de collecte des documents. La notion de numérique n'étant que rarement liée à celle d'archives dans l'esprit des personnes auprès de qui s'effectue cette collecte, il convient de se montrer très vigilant et de faire preuve de beaucoup de pédagogie afin d'éviter un traitement différent du papier et du numérique qui en droit français n'existe pas : tous les documents produits par les administrations sont archives publiques, quelle que soit leur nature.

Pour résumer, les avantages sont indéniablement techniques, et les inconvénients sont en quelque sorte philosophiques, mais avec des implications très concrètes.

4) Quels moyens votre service met-il en œuvre concernant l'accès numérique ?

Droits d'auteur réservés. OU (cc) BY-NC-ND

Les Archives de France sont un service interministériel qui fait fonction de tête du réseau des archives publiques. C'est notamment le service qui pilote le projet de portail Francearchives.fr, qui a vocation à servir de point d'entrée à l'ensemble des archives numérisées et inventaires produits par les archives en France. Ce site regroupe également un certain nombre d'informations mises à la disposition du public, par exemple les données produites dans le cadre de l'enquête statistique annuelle, dont le millésime 2016 vient d'être publié en ligne.
Les Archives de France publient également certains jeux de données sur data.gouv.fr, notamment ces chiffres de l'enquête annuelle.

5) Quel accès dans la tension entre mémoire et droit à l'oubli ?

L'accès simplifié permis par le numérique entraine comme je l'ai dit un changement d'échelle : on peut consulter en quelques minutes des dizaines de dossiers, pour lesquels il aurait auparavant fallu se déplacer et attendre communication en salle. Or, cet accès simplifié et à grande échelle correspond à une vraie demande sociale : on voit notamment dans le cadre de projets comme le Grand mémorial un vrai désir d'avoir accès aux informations relatives aux soldats de la première guerre mondiale et d'écrire l'histoire de sa famille.

Pourtant, vous avez peut-être connaissance des débats qui ont eu lieu autour du règlement européen sur les données à caractère personnel (DCP) : de peur que les données de tout un chacun puissent être utilisées sans contrôle, ce texte prévoyait dans une première version la destruction des DCP collectées par les administrations à l'issue de leur durée d'utilité administrative (je simplifie bien entendu le débat). Il a finalement été possible d'insérer une exception pour les DCP destinées à une conservation définitive.

Ce débat est selon moi tout à fait représentatif du problème que vous évoquez : actuellement, nous voulons tout savoir sur tout le monde, mais nous souhaitons protéger au mieux nos propres données et pouvoir les détruire lorsque nous n'en avons plus besoin, ce qui est assez paradoxal quand nous sommes ravis de trouver les mêmes informations sur nos ancêtres.

La tension principale vient en réalité de l'absence de profondeur historique donnée à la réflexion : les règles qui valent pour les informations produites aujourd'hui cessent au bout d'un moment d'être pertinentes, notamment lorsque la personne concernée est décédée.

Droits d'auteur réservés. OU (cc) BY-NC-ND

Cela explique la mise en place de délais plus longs pour la consultation sur internet que pour la consultation en salle pour les documents comportant des données à caractère personnel (les documents comportant des secrets de vie privée, communicables à 50 ans, ne sont diffusables sur internet par les services d'archives qu'au bout de 100 ans, voire 120 ans pour certains secrets considérés comme plus sensible comme l'intimité de la vie sexuelle, la religion ou les opinions politiques). Cela garantit la protection au mieux des informations contenues dans les documents, qui ne sont mises en lignes qu'après le décès des personnes concernées.

6) Comment voyez-vous l'accès de demain (perspectives d'évolution) ?

Les règles de mise en ligne que j'évoque juste au-dessus sont appelées à évoluer : les usagers souhaitent en réalité la diffusion des documents qui concernent leur famille afin de pouvoir faire leurs recherches plus aisément. Seule une demande de retrait a eu lieu depuis le lancement du Grand mémorial : il y a donc une acceptation sociale de la diffusion des informations relatives aux grands-parents et au-delà, lorsqu'ils sont décédés.

Il serait souhaitable que l'accès soit encore simplifié, en permettant un accès distant généralisé. Je ne pense pas que cela entraîne une disparition des salles de lectures, et il ne s'agirait même dans ce cas que d'une rematérialisation d'une salle de lecture distante, et non d'une disparition. Les projets permis par le numérique, notamment en termes de valorisation, sont également appelés à se développer.

Droits d'auteur réservés. OU (cc) BY-NC-ND

ANNEXE 4 : ENQUÊTE AUPRÈS DE L'ÉQUIPE VITAM

1) Quelle est votre définition de l'accès aux archives publiques ?

Personne A :

Capacité des citoyens et autres utilisateurs à rechercher et consulter l'information produite par les administrations ou leurs délégataires dans l'exercice de leurs missions de service public, dans le respect du droit d'en connaître.

Personne B :

L'accès aux archives publiques s'entend de deux façons pour un usager :

- Le fait d'aller dans une salle de lecture, physiquement ou virtuellement, en vue de consulter des documents d'archives publiques,
- La consultation de documents d'archives publiques, librement communicables (ou après avoir fait une demande de dérogation).

Pour un archiviste, on pourrait ajouter la gestion de la communicabilité, qui est un préalable à l'accès.

Personne C :

L'accès aux archives publiques est la recherche et la communication des documents et données conservées par les services publics d'archives aux utilisateurs (chercheurs citoyens ou autres), afin de répondre à leurs besoins, dans le cadre des délais de communication. C'est le fondement de la conservation des archives : constituer un patrimoine qui doit être consulté et exploitable.

2) En quoi le numérique est-il venu modifier cet accès ?

Personne A :

Il permet de rendre plus aisément le service attendu.

Personne B :

L'accès aux archives nécessitait auparavant pour l'usager un ou plusieurs déplacements physiques aux archives ou une demande par correspondance d'un document qui n'aboutissait pas forcément. Avec le numérique, le déplacement physique n'est plus forcément nécessaire ou, du moins, il peut être limité. Les

Droits d'auteur réservés. OU (cc) BY-NC-ND

inventaires sont en ligne et permettent de cibler ses recherches et d'optimiser son déplacement dans un service d'archives. Ex :

- Réservation en ligne évitant d'avoir à attendre les levées le jour de sa venue aux archives.
- Lister avec les outils en ligne les archives à consulter, plutôt que de le faire, le jour de sa venue aux archives, sur des instruments de recherche papier non consultables en ligne, comme cela était le cas ultérieurement.

Avec le numérique, on peut également accéder directement de chez soi à des archives, librement communicables. On a désormais accès à des documents de tous pays, sans difficultés, sans avoir à se déplacer. On peut également communiquer en ligne avec certains services d'archives et faire des demandes en ligne de reproduction d'archives (téléservices - ex : AD de la Haute-Loire, je crois).

Personne C :

Le numérique tend à faciliter l'accès en n'obligeant pas forcément le lecteur à se déplacer géographiquement, mais cela pose la question de la sécurité des systèmes d'informations : difficile faire circuler des archives non communicables sur les réseaux.

3) Quels sont les avantages et les inconvénients de l'accès via le numérique ?

Avantages - personne A :

- Possibilité de dissocier lieu de conservation des informations et lieu de diffusion.
- Plus systématiquement besoin de se déplacer dans un service d'archives et de recourir à un archiviste pour assurer la médiation.
- Nouveaux usages permis.
- Possibilité de dédoubler l'information, à savoir que l'information peut rester chez le producteur tout en étant déjà prise en charge par les services d'archives pour remplir d'autres usages.
- Richesse des requêtes possibles.

Avantages - personne B :

Droits d'auteur réservés. OU (cc) BY-NC-ND

- Baisse des déplacements et des coûts de voyage
- Consultation directe en ligne de documents d'archives et d'instruments de recherche.

Avantages - personne C :

Le numérique ouvre certainement des potentialités d'exploitation des archives beaucoup plus larges. Les archivistes doivent réfléchir aux progrès que le numérique pourrait apporter aux utilisateurs : recherches plus simples, duplication avec leurs propres annotations, anonymisation de corpus pour exploitation historique avant fin de DUA.

Inconvénients – personne A :

- Complexité juridique croissante.
- Un marché de l'accès à l'information de plus en plus concurrentiel (bibliothèques, archives, documentation, open data, etc)
- Besoin de contextualisation semble moins évident mais encore plus nécessaire.
- Difficulté à faire comprendre que la qualité des données peut être variable.

Inconvénients – personne B :

- L'usager n'a accès en ligne qu'à certains corpus (tout n'est pas accessible). Il peut manquer de visibilité sur des fonds complémentaires, non visibles et accessibles en ligne. Ce qui peut finalement le faire passer à côté de certains types d'archives.
- L'assistance en ligne, les téléservices ne sont pas développés uniformément. Les politiques d'accès aux ressources en ligne divergent.
- Plusieurs lois se recoupent sur l'accès en ligne : code du patrimoine, loi CNIL et loi CADA (et leurs pendants européens), droit d'auteur, de la propriété intellectuelle. Un document non accessible en ligne peut l'être en salle de lecture. Ce n'est pas toujours clairement connu d'un usager.

Inconvénients – personne C :

Le numérique pose la question de la visualisation et de la forme des documents : comment remonter une base de données pour consultation, les archives conservées n'auront plus forcément leur forme d'origine après opération de pérennisation…

Droits d'auteur réservés. OU (cc) BY-NC-ND

4) Quels moyens votre service met-il en œuvre concernant l'accès numérique ?

Personne A :

Conception et réalisation d'une solution logicielle d'archivage numérique :

- Utilisant les technologies de big data pour permettre des recherches transverses.
- Interfaçable avec tout type d'applications via des API (logique état plate-forme) et donc agnostique du client final (possibilité de fournir des services aussi bien à un outil d'archivistes comme de bibliothécaires, de documentalistes d'open data ou autre).

Personne B :

Une partie des fonctionnalités concerne l'accès. Il s'agit de permettre l'accès aux archives en fonction des règles de gestion qui leur sont associées (règles de communicabilité, règles de classification, règles de réutilisation, règles de diffusion), et des utilisateurs (internautes, services producteurs). En fonction des droits de ces derniers, ils auront accès aux archives numériques ou à une version de consultation.

Personne C :

Vitam est un logiciel de back office. On cherche, par la conception du logiciel, à trouver les meilleurs moyens de conserver les archives numériques tout en garantissant et facilitant l'accès. On travaille sur les modèles de données et arborescence dans cette optique ; Les projets d'implémentation ministériels (Archives nationales, SHD, archives diplomatiques…) sont ensuite chargés de créer les front offices qui vont rendre l'accès facile et attractif pour leurs utilisateurs.

5) Quel accès dans la tension entre mémoire et droit à l'oubli ?

Personne A :

Difficile d'apporter une réponse :

- La donnée personnelle est un des produits phare des services d'archives.
- Les services d'archives permettent de garantir le droit à l'oubli dans le contexte de production tout en préservant, par l'archivage, la mémoire de l'information et sa réutilisation sur le long terme, dans le respect de la vie privée.

Droits d'auteur réservés. OU (cc) BY-NC-ND

- En même temps, il est difficile de se projeter vu la richesse des réutilisations possibles et la dangerosité de celles-ci pour la vie privée des personnes.
- Le numérique offre des possibilités d'anonymisation non encore exploitées.

Dans l'immédiat, je dirais que les services d'archives ont un rôle à jouer dans la sensibilisation aux problématiques d'accès aux données personnelles.

Personne B :

Il faut un accès maîtrisé, tel qu'il existe actuellement : on ne rend accessible en ligne que des archives à données personnelles avec des délais de communicabilité longs, voire on ne les rend accessible que via des plateformes sécurisées, non indexées par les moteurs de recherche externes. Ce sont ces derniers qui posent problèmes concernant le droit à l'oubli.

Personne C :

Ces questions sont régies par les délais de diffusion des données sur internet qui protègent l'individu. Les services d'archives peuvent garantir la conservation d'information utiles sans les divulguer pendant les délais de non-diffusion ou de non communicabilité. À mon sens, le droit à l'oubli est important en ce qui concerne des délais courts et des diffusions très larges d'informations sur Internet (il est plus critique en ce qui concerne la publication d'information sur les réseaux sociaux). Pourrait-on se permettre d'anonymiser des corpus pour les diffuser plus rapidement aux chercheurs sans prendre de risque ? C'est à mon sens une réflexion à mener afin de valoriser plus rapidement les fonds conservés.

6) Comment voyez-vous l'accès de demain (perspectives d'évolution) ?

Personne A :

De plus en plus dissocié de la conservation et de sa localisation, et surtout via des canaux de plus en plus multiples (points comme fournisseurs d'accès).

Personne B :

- Un développement des téléservices en ligne.
- Une meilleure formation des internautes sur les outils disponibles en ligne et en interne (au sein des services d'archives), sur les différents droits couvrant les données (personnelles) mises en ligne.

Droits d'auteur réservés. OU (cc) BY-NC-ND

- Une simplification des droits sur l'accès.

Personne C :

- Développer de nouveaux services aux utilisateurs : bureau et outils numériques.
- Utiliser le réseau existant de services d'archives pour la communication des archives sous dérogation : transmission sécurisée des archives entre services et accueil de l'usager indépendamment du lieu de production/conservation des archives.
- Développer un ensemble de services à distance pour les usagers : hotline, chat… conseils à distance mais simultané comme on peut le faire en salle de lecture.
- Les archivistes doivent réfléchir à la recherche d'information : faut-il faire des instruments de recherche, la constitution de bases de données uniques est-elle réaliste, comment dans ce cas permettre la contextualisation qu'offrent les IR ? Les travaux Vitam ont montré que les utilisateurs attendent un mode de recherche simple type « Google » : un champ permettant de saisir des mots-clés.

Droits d'auteur réservés. OU (cc) BY-NC-ND

ANNEXE 5 : ENQUÊTE AUPRÈS DE CYRIL LONGIN

1) Quelle est votre définition de l'accès aux archives publiques ?

Elle est réglementaire. Deux éléments. 7 Messidor An II, le citoyen a accès aux archives de manière libre et gratuite. Notion fondamentale sur laquelle le reste est fondé. Ensuite c'est du technique, pratique, juridique. 2008, après le terme d'accès, l'usager sollicite, fait la demande. Les gens veulent avoir accès à l'information. On a changé l'approche. La loi CADA c'était faciliter les relations entre l'administration et les usagers. Aujourd'hui le paradigme a changé, avec une politique plus de diffusion, de mise à disposition que dans le sens inverse.

Ma définition, la mise en ligne n'est en aucun cas une obligation, sauf pour ce qui est nativement numérique. Ce n'est pas encore écrit dans la loi.

2) En quoi le numérique est-il venu modifier cet accès ?

C'est internet. Tout commence par internet. Il parait logique de mettre la numérisation sur internet. La numérisation c'est le même processus que le microfilmage, on sauvegarde l'original. On pourrait toujours faire du microfilmage. L'intérêt de la numérisation, c'est qu'elle permet la diffusion. Aujourd'hui, je ne connais plus de services d'archives qui continuent de microfilmer. C'est la possibilité de mettre à disposition les formats de substitution. On ne numérise plus seulement pour conserver l'original, mais pour mettre en ligne. Aujourd'hui on reprend les originaux pour la lisibilité sur internet. L'original est déjà conservé. Aujourd'hui, est ce que la mise en ligne est une suite logique à l'accès, c'est une question. Le public l'entend comme ça mais la loi ne le dit pas. On est resté sur une approche 78.

3) Quels sont les avantages et les inconvénients de l'accès via le numérique ?

7 jours sur 7, 24h sur 24, possibilité de réutiliser. La numérisation permet de donner accès à un plus grand nombre. Donner accès à des documents compliqués à communiquer en salle de lecture, archives orales, négatifs, plans de 17 mètres de long. Donne accès tout court. Plaque de verre.

Détails qu'on ne voit pas à l'œil nu. L'inconvénient en termes d'accès, c'est que les gens ne se rendent pas compte du travail en amont. Un peu comme l'open data. C'est une évidence. Le deuxième point négatif, c'est la question des données à caractère personnel et de ce qui n'est pas communicable. Comment donner accès

Droits d'auteur réservés. OU (cc) BY-NC-ND

à des choses non diffusable en ligne ? Accès restreint, par dérogation. On va devoir réfléchir aux accès sécurisés, adresse IP, comptes privés ? Il faut se poser la la question de l'accès restreint.

4) Quels moyens votre service met-il en œuvre concernant l'accès numérique ?

Le poste est vacant. Service assez avancé. Le premier en 2006 à avoir un site propre. À la pointe. Un peu moins après. On a changé l'outil au bout de 8 ans. En moyens humains c'est 1,5 à 2 ETP. Avec Orléans, on est quand même très avancés par rapport à d'autres. Site participatif, 300 articles sur l'histoire de la ville, blog, information géographique, on va ouvrir une plateforme d'archives orales. Ressources numériques, bases de données, textes historiques. En termes de moyens finances, pour la numérisation, 20 000 à 25 000 euros par an.

5) Quel accès mettre en place dans un contexte de tension entre mémoire et droit à l'oubli ?

Les archivistes sont en dehors de ça. Défense des intérêts. Contraint par l'administration. En dessous de 100 ans, les données à caractère personnel on ne les met pas. Le droit à l'oubli n'a plus de sens. Pour la mise à dispo des ressources sur internet, la question ne se pose pas puisque qu'il y a la législation. Les archivistes sont là pour permettre aux historiens et particuliers de conserver leur mémoire collective et individuelle tout en garantissant le droit individuel des personnes. Tout en sécurisant la question des accès. Il y a un droit à l'oubli, mais aussi un droit à l'histoire. Les deux ne sont pas contradictoires. C'est une question de gouvernance de l'information. Le problème il est là. Les archivistes sont là pour être garants. Ils ont une certaine indépendance et distance par rapport à l'information qu'un particulier ou journaliste n'aura pas.

6) Comment voyez-vous l'accès de demain (perspectives d'évolution ?)

Accès sécurisé, individuel. Mise à disposition des ressources qu'on conserve et qu'on produit également. On produit de la donnée. La question des sites d'archives est à mettre en lien avec les sites d'Open Data. Nos sites internet, on les voit bien assez proches des plateformes open data. On va devenir le réceptacle de données produites. Entrepôt de données. L'approche de l'accès sera inversée. Ce sera une obligation de diffuser.

Droits d'auteur réservés. OU (cc) BY-NC-ND

TABLE DES MATIÈRES

SIGLES ET ABRÉVIATIONS 9

INTRODUCTION 11

CHAPITRE I : LA MISSION TRADITIONNELLE DE COMMUNICATION DES ARCHIVES PUBLIQUES 15

I> historique de l'accès aux archives 15

A> La loi du 7 messidor an II 15

B> Du XIXème siècle au 3 janvier 1979 16

II> La règlementation du droit d'accès 18

A> Les délais de communicabilité 18

B> Les accès particuliers 22

1) L'accès privilégié 22

2) L'accès professionnel 23

3) L'accès par dérogation 24

C> Le rôle de la CADA 26

1) Historique et missions 26

2) Le rapport d'activité 2015 27

a) Une augmentation significative du nombre de saisines en 2015 28

b) La CADA et l'évolution du contexte législatif 31

III> les modalités d'accès matériel aux archives publiques, dans le contexte papier 31

A> Les dispositions législatives 32

B> La salle de lecture 32

C> L'aide à la recherche 33

D> Les chiffres de 2015 34

CHAPITRE II : LA RÉVOLUTION NUMÉRIQUE OU LE PASSAGE DE LA COMMUNICATION À LA DIFFUSION 37

I> Un nouveau contexte – tant législatif que sociétal – favorable à l'ouverture des données publiques 37

A> L'ouverture des données 38

1) Le mouvement open data 38

a) Principes 38

b) Objectifs 38

2) Historique du mouvement 40

a) Les anglo-saxons à la barre 40

Droits d'auteur réservés. OU (cc) BY-NC-ND

b) Une initiative poursuivie en France 41
c) Un mouvement soutenu et encouragé par des associations 43
B> Évolutions législatives et règlementaires 44
1) Le plan d'action national pour une action publique transparente et collaborative 44
2) La directive européenne PSI et la loi Valter 46
3) La loi Lemaire, dite « pour une République numérique » 47
a) L'ouverture des données par défaut 48
b) La création d'un service public de la donnée 51
c) L'ouverture des données d'intérêt général 51
d) Le libre accès aux travaux de la recherche publique et l'exploration des données 52
II> Les nouvelles formes d'accès 52
A> Les portails open data 52
1) Enjeux, acteurs et licences 53
a) Les enjeux, pourquoi ouvrir les données ? 53
b) Les acteurs, avec qui partager les données ? 53
c) Les licences, sous quelles conditions partager les données ? .. 54
2) Construction des portails 54
3) Les jeux de données 57
a) Les approches 57
b) Les typologies 58
c) Les principales thématiques 59
d) Une absence d'homogénéité 59
B> Les services publics d'archives et le numérique 60
1) Les atouts du numérique pour les services publics d'archives 61
a) Pourquoi mettre en ligne ? 61
b) Pour répondre à quels usages ? 63
c) Quels droits d'accès et d'utilisation ? 63
2) Les sites web des services publics d'archives 64
a) Stratégie web, navigation et ergonomie 64
b) Moteur de recherche 65
c) Portails et moissonnage des données 66
d) Quelques données chiffrées 67
3) Les archives numérisées et les outils disponibles en ligne 68
a) Richesse et diversité des fonds numérisés 69
b) Conditions d'accès et de recherche 70
III> Vers de nouveaux acteurs ? 74

Droits d'auteur réservés. OU (cc) BY-NC-ND

A> Le cas d'un EPIC : la SNCF 75

B> Le cas d'une entreprise privée : Filae 81

CHAPITRE III : UN ACCÈS TRANSFORMÉ 85

I> L'accès autour des données à caractère personnel 85

A> L'autorisation unique 029 et la règlementation européenne 86

B> Le déréférencement 89

1) L'arrêt Google Spain 90

2) Une législation encore à définir 92

C> La tension entre mémoire et droit à l'oubli 93

II> De nouvelles pratiques d'accès 96

A> L'impact du numérique sur la méthodologie de recherche archivistique 96

B> La réutilisation 100

C> Des services innovants : le cas des AD des Hautes-Alpes 103

1) Un public devenu virtuel 104

2) Un renversement du pouvoir 104

3) Repenser le modèle d'accès et de recherche aux archives publiques 105

CONCLUSION 111

SOURCES 115

BIBLIOGRAPHIE 117

ANNEXES 121

TABLE DES MATIÈRES 139

Droits d'auteur réservés. OU (cc) BY-NC-ND