

Les archives dans une société numérique

Luciana Duranti et Corinne Rogers

Introduction

Ce texte porte sur les responsabilités du système archivistique canadien envers le résidu d'origine numérique des activités de la société et les stratégies pour sa préservation. Il décrira d'abord les caractéristiques des sous-produits numériques de l'activité des organisations et des individus, décrira ensuite les responsabilités qui en découlent pour le système archivistique canadien et envisagera la manière dont il pourrait répondre aux attentes de la société canadienne.

Les caractéristiques de l'information d'origine numérique d'aujourd'hui (données, documents, dossiers)

Comprendre les caractéristiques de l'information d'origine numérique d'aujourd'hui n'est pas possible sans d'abord analyser le comportement et les attentes de ses créateurs et utilisateurs.

Au cours des dernières années, un fossé générationnel est clairement apparu entre ceux qui sont nés avant et après 1981, ce dernier groupe étant défini comme la génération du millénaire, ou la génération Y. La tendance de l'ancien groupe a été de transposer à l'environnement numérique le même comportement que pour l'analogique, en se fondant sur les mêmes valeurs et convictions, notamment en ce qui concerne les concepts fondamentaux de l'accès, la vie privée, les droits de propriété intellectuelle, la propriété, la création, la conservation, la responsabilité, la reddition de comptes, la mémoire enregistrée, et la préservation. En conséquence, l'analyse et la recherche rigoureuses des produits numériques a abouti à des pratiques de tenue de dossiers et de préservation des documents largement acceptées, des normes internationales évolutives et applicables, et, pour le Canada, à une génération d'archivistes confiants dans leurs connaissances et dans leur capacité à acquérir et à maintenir le contrôle sur la production actuelle de notre société numérique, au moins aussi longtemps que les ressources technologiques et financières suffisantes seront disponibles. En d'autres termes, les sous-produits documentaires numériques des activités de la génération pré-millénaire, dans les contextes organisationnels et personnels, peuvent être conservés et rendus accessibles sans la nécessité de changements fondamentaux dans le système archivistique canadien.

Ce n'est pas le cas lorsque l'on considère le comportement numérique et les attentes de la génération Y, qui a maintenant atteint l'âge adulte et a commencé sa vie professionnelle dans un environnement caractérisé par la convergence des médias et la connectivité

constante, souvent décrit comme mobile et liquide. Il s'agit d'un environnement dans lequel les anciennes et les nouvelles technologies se confondent dans le même appareil de sorte que les caméras, les téléphones, les ordinateurs et l'Internet sont parfaitement intégrés. Le réseautage social et les sites de partage de photos encouragent les utilisateurs à communiquer en ligne et à donner à leurs contacts la permission d'ajouter des commentaires, des notes, des étiquettes, des images, etc., de sorte que les frontières entre les producteurs de contenu et les utilisateurs deviennent floues. Les sous-produits de cette interaction ne sont plus des entités finies, mais des processus qui sont toujours en train de changer, véhiculés par la plate-forme technologique qui les accueille grâce à des métadonnées, des protocoles et des codes. Ce phénomène est bien connu des chercheurs en sciences sociales, qui identifient ses participants, qu'il s'agisse d'organisations ou d'individus, en utilisant le néologisme de *produsateur*, un terme qui confond volontairement les rôles de producteur et d'utilisateur.

Ainsi, on peut voir que la pratique relativement nouvelle de BYOD, ou Bring Your Own Device (Apportez votre propre appareil), dans les organisations et les entreprises constitue un problème non seulement parce que l'information numérique de l'entreprise n'est plus capturée ou déplacée dans un système d'archivage des dossiers central, mais aussi parce qu'elle est rarement complète, manque de fixité, est difficile à attribuer à un auteur et/ou propriétaire, surtout en raison de l'esprit de collaboration avec lequel les gens agissent, et n'est pas clairement liée à un processus, une action ou une opération d'affaires. Il est très difficile dans ce cas d'attribuer la responsabilité ou de demander des comptes. La situation présente une plus grande complexité lorsque les employés, ainsi que leurs appareils, utilisent leur propre « cloud » (plateforme en nuage), ou plusieurs « cloud » différents dans lesquels leurs résultats sont stockés de façon généralisée, la plupart liés à des applications plutôt qu'à des fonctions de l'entreprise (par exemple, DropBox ou Google Docs). « Un récent sondage indique que plus de la moitié des travailleurs de la génération Y sont prêts à aller à l'encontre des politiques de l'entreprise sur l'utilisation de leurs propres appareils ou de plateforme en nuage si celles-ci posaient problème à leurs activités personnelle et professionnelle et réseaux sociaux. »¹ Le travail sera mené de la manière la plus efficace, indépendamment du fait qu'il laisse une trace - données ou documents - pouvant servir de référence pour d'autres actions ou à titre de preuve des actions et des transactions enregistrées.

¹ http://www.itworldcanada.com/article/gen-y-workers-ok-with-flouting-cloud-byod-policies/86041?sub=67871&utm_source=67871&utm_medium=comminfra&utm_campaign=enews. Pour les rapports sur d'autres sondages, visitez le <http://www.ashridge.org.uk/website/content.nsf/wFARCRED/Culture+shock+Gen+Y+and+their+managers+around+the+world> et <http://www.pwc.com/gx/en/managing-tomorrows-people/future-of-work/key-findings.jhtml>.

Au fil du temps l'information numérique *produit* par ces employés augmente en volume, et, étant répartie entre les différents appareils et plates-formes multiples de réseautage social en ligne, devient ingérable, et ses producteurs choisissent tout simplement de la laisser s'accumuler, étant donné la disponibilité de grandes quantités de stockage gratuit sur les plateformes en nuage. Ainsi, le problème de l'absence de documents utilisables comme précédent ou comme preuve est aggravée par l'abondance incontrôlée, une question qui commence à ennuyer aussi les *producteurs*. Mais la technologie a déjà offert une solution par la suppression automatique après un certain laps de temps, un laps de temps qui souvent n'est pas conforme au programme de conservation et d'élimination approuvé par l'employeur, mais se fonde plutôt sur l'intention de la communication qui génère l'information, telle qu'elle est définie par la personne à l'origine. Pourquoi un employé n'envisagerait-il pas la possibilité que l'information en question puisse être nécessaire plus tard ou même à long terme ? Tout simplement parce que la génération Y considère que la communication est plus importante que la mémoire et que le matériel qu'elle génère est censé avoir un impact immédiat et être consommé instantanément. Ainsi, il n'y a pas d'attentes de conservation, seulement le désir de générer des données tout au long de chaque événement, en continu, comme si rien ne pouvait arriver à moins d'être communiqué, mais sa présence n'a pas été supprimée par la disparition de ses traces numériques.

Si, pour un moment, nous regardions à travers la lentille des archives, il est assez clair que ce comportement dans les organisations sape non seulement leur productivité et croissance à long terme, la conformité, la responsabilité, la responsabilisation, et la capacité de se protéger, au présent et dans le futur, ainsi que la construction d'une culture organisationnelle, mais aussi toute chance pour les archives d'identifier des documents, des séries, ou des fonds, d'effectuer une évaluation systématique, et d'acquérir les sources de la mémoire de la société. Et la situation est pire si, encore une fois en tant qu'archivistes, nous pensons aux fonds des associations volontaires, des ONG et des individus. Alors que dans tous ces cas, la plupart des activités sont menées dans plusieurs appareils et environnements en ligne, les individus ajoutent un niveau de complexité à leur production numérique en sélectionnant des langues de communication qui sont externes aux paramètres établis, par exemple en ayant de plus en plus recours à des images et à l'abandon des mots (par exemple, l'envoi de la photo d'un bâtiment et d'une carte Google Map au lieu d'une invitation à se rendre à un endroit et une adresse) . En outre, pour les archivistes, traquer les composantes d'une archive personnelle au fil du temps impliquerait d'avoir accès aux sites de réseautage social du créateur, ses comptes de téléphonie mobile, ses comptes de messagerie Twitter et de sites de partage de photos, etc. Même si le

matériel existait encore et pourrait être situé et accessible, il serait encore difficile d'identifier les « fonds personnels. »²

Les caractéristiques de l'information numérique *produit* par la génération Y sont donc tout à fait claires : lorsqu'elle est traçable et qu'elle possède des limites identifiables, il s'agit d'une entité co-créée (par des personnes et des plates-formes, des codes logiciels et des protocoles des applications hôtes), toujours en devenir, générée pour servir à des fins multiples dans de multiples contextes, composée de plus d'un support, réalisée, plutôt que livrée, et donc éphémère dans sa nature, généralement accessible par tous ceux qui partagent la même plate-forme et / ou appartiennent aux mêmes cercles, et protégée par des droits intellectuels.

Comment ce brouillage des frontières entre public et privé, entre organisationnel et personnel, entre éphémère et permanent, entre complet et dans un état de devenir, entre générée par une plate-forme et générée par les utilisateurs, entre privée et ouverte, entre enregistrée et réalisée, etc., répond-il aux attentes de ceux qui se considèrent eux-mêmes comme des utilisateurs de documents et d'archives? Leur comportement a-t-il changé aussi ?

Les utilisateurs de documents et des archives appartiennent encore pour la plupart à la génération pré-1981, ce qui signifie qu'ils s'attendent à ce que l'environnement numérique facilite les comportements déjà établis, mais qu'ils découvrent aussi les nouvelles possibilités offertes par les technologies complexes. Les chercheurs qui ont toujours travaillé en collaboration s'attendent maintenant à le faire dans un environnement virtuel, les utilisateurs de matériel dans les fonds d'archives s'attendent à ce que ces documents soient mis à disposition sur l'Internet, instantanément et dans les formats requis et interopérables ; les citoyens intéressés par les actions du gouvernement et des sociétés s'attendent à avoir un libre accès aux documents de ces actions dans le temps ; les scientifiques du monde des données et des statistiques sont intéressés aux initiatives de données ouvertes et s'attendent à la traçabilité de ces données ; et un nombre croissant d'utilisateurs du secteur privé et d'universitaires utilisent diverses sources de données volumineuses pour créer des *mashups* de données pour une variété de fins. La communauté archivistique canadienne connaît depuis longtemps les comportements et les attentes de ces utilisateurs, mais le changement le plus notable est l'intérêt croissant pour les données, par opposition aux documents, et pour les grands regroupements de données de recherche, de données d'observation, et de données recueillies par le gouvernement pour soutenir la prise de décision. La considération clé ici est que cet intérêt pour les données comporte aussi un fort intérêt pour l'exactitude, la fiabilité et la traçabilité de sources authentiques. Cette exigence de fiabilité, facteur important dans le choix d'utiliser des sources primaires,

² Pour plusieurs de ces idées sur le comportement individuel, nous sommes reconnaissants à la doctorante de UBC Jessica Bushey, dont la recherche les a découverts dans la littérature des sciences sociales.

devient une demande explicite pour une chaîne ininterrompue documentée de conservation responsable : en d'autres termes, la confiance traditionnellement accordée au conservateur requiert maintenant aussi des motifs transparents.

Les responsabilités du Système archivistique canadien - reconnaître les défis

Comment les institutions d'archives, les organisations, les unités et les programmes - le système archivistique canadien – peuvent-ils aborder les opportunités et les défis posés par l'évolution des façons dont les Canadiens communiquent et travaillent ? L'orientation de la communauté archivistique sur la « mémoire » et les « preuves » encourage la (mauvaise) perception populaire des archives comme s'occupant uniquement des documents du passé (peu importe leur forme). Cela tend à marginaliser les archives et les archivistes et, par le fait même, on perpétue le mythe du sauveur passif (même si nous savons que ce mythe est tout à fait faux). Nous savons depuis longtemps que notre matériel d'origine numérique ne survivra pas dans une forme utilisable et fiable si sa conservation n'est pas considérée dès le moment de sa création. Bien que ce soit déjà assez difficile pour les documents d'origine numérique qui sont une réflexion de leurs homologues traditionnels, les documents créés en collaboration et en ligne créent de nouveaux défis. Identifier les questions de création, de propriété, l'accès et les droits de propriété intellectuelle ne sont que quelques exemples du défi qui est maintenant devant nous. Quelles sont les responsabilités et les stratégies du système archivistique canadien à soutenir une création contrôlée des données, des documents et des dossiers qui puisse démontrer des processus transparents et une chaîne de garde ou de provenance, et dont la précision, la fiabilité et l'authenticité peuvent être assurées au fil du temps et au travers des plates-formes technologiques? Maintenant plus que jamais, les archivistes doivent se positionner sur le côté des créateurs s'ils veulent être en mesure de remplir leur mandat d'organisation : ils ont besoin d'identifier les documents et les données dignes de capture et la conservation à long terme lors de la création, de déterminer leur capacité à préserver, et à contrôler leur utilisation et leur transformation à travers le temps. Cela signifie comprendre l'évolution des frontières entre public et privé, organisationnel et personnel, auteur et propriétaire, créateur et conservateur, producteur et utilisateur, et les attributs du *produsateur*.

Un cadre réglementaire général doit être élaboré et mis en œuvre pour garantir et protéger la création et la gestion de ces nouveaux documents organisationnels, pour s'adapter à la nouvelle réalité des politiques BYOD, et l'intégration de l'information personnelle et organisationnelle, privée et publique qui résulte inévitablement de l'adoption d'un tel cadre. En son sein, des politiques spécifiques devraient exiger que des calendriers de conservation et d'élimination soient mis en place pour les documents créés dans des appareils personnels ainsi que dans les systèmes de gestion des documents de l'organisation, et que les employés ne puissent pas appliquer de dates de suppression pour

les documents désignés pour la conservation à long terme. Il faut également envisager que la connectivité de la base de données du réseau transforme l'expérience de l'utilisateur et pourrait permettre à des individus possédant des aptitudes techniques d'accéder à des informations qui doivent être protégées en vertu des lois de la vie privée, et on doit s'assurer que cela ne se produise pas.

Dans le domaine des archives personnelles, le système archivistique canadien a besoin de comprendre et de déterminer comment identifier les archives d'une personne sur le long terme (par exemple, comment retracer la correspondance d'une personne avec un groupe d'amis, au fil du temps, à travers les comptes de partage de photos, les sites de réseautage social, les comptes de téléphonie mobile, les comptes de messagerie, Twitter, etc.), pour localiser ce qui existe encore, pour y avoir accès (que faire avec la protection par mot de passe ?), pour le récupérer, et donner un sens à une surabondance de l'éphémère, puis fournir des conseils aux particuliers et aux petites communautés sur la façon d'apprécier, de maintenir et de protéger une mémoire documentaire.

Réalisation : les enjeux de l'avenir

La société numérique est dans un état continu de devenir. Plus que jamais, le système archivistique canadien doit prendre la responsabilité des fonctions d'archives qui sont à sa base (évaluation et acquisition, aménagement et description, conservation et préservation, gestion et administration, référence et accès), mais il doit le faire à travers l'ensemble du cycle de création, de gestion, de l'utilisation et de la réutilisation, dans les secteurs public et privé, d'une manière intégrée.

Les fonctions de base des archives ne peuvent être menées à bien que lorsqu'elles sont fondées sur une base solide de collaboration interdisciplinaire entre la recherche et l'éducation. Les archivistes ne sont pas des informaticiens, et ne devraient pas aspirer à l'être non plus. Nous n'influencerons pas l'orientation de l'évolution technologique, mais nous ne pouvons pas simplement être des observateurs. Nous devons tenter de prédire cette direction et ces développements, et d'y être sensibles si nous voulons identifier, capturer et préserver le patrimoine de demain. Cela n'arrivera pas en s'asseyant et en attendant que le milieu universitaire étudie le problème et propose des solutions. Tous les archivistes peuvent et doivent être parmi ceux qui conçoivent des solutions. Il faut travailler ensemble dans les institutions, les dépositaires, les départements universitaires, avec des collègues dans une variété d'autres domaines et disciplines si nous voulons répondre de façon positive aux questions qui ont été posées par le rapport intérimaire de CAST.³

³ Le rapport intérimaire de CAST pose les questions suivantes: Avec l'utilisation des nouvelles technologies de communication, quels types de documents sont créés et comment les décisions sont-elles documentées?

Un modèle de ce type d'activité interdisciplinaire existe à l'Université de la Colombie-Britannique. Le InterPARES Trust est un partenariat de recherche international impliquant des archivistes, des gestionnaires de documents, des bibliothécaires, ainsi que des avocats, des informaticiens, des spécialistes de l'investigation numérique, des experts en sécurité de l'information, des journalistes, des éthiciens, des spécialistes de la politique, et plus encore, de l'industrie, du gouvernement et du milieu universitaire. Les experts de chaque discipline écoutent et apprennent des autres ; la force de plusieurs disciplines converge vers un seul objectif – trouver des moyens nous permettant de faire confiance aux données, documents et enregistrements qui se trouvent dans les environnements en ligne, et les implémenter.

Les connaissances acquises par cette collaboration et d'autres similaires permet certes à toutes les organisations participantes d'être bien préparées pour l'émergence de nouvelles questions, mais tous ces efforts seraient perdus si les connaissances acquises ne sont pas transmises aux nouvelles générations de professionnels. Ceci doit être fait en impliquant les étudiants dans la recherche collaborative, qu'il s'agisse d'environnements d'archives ou de créateurs. La recherche interdisciplinaire doit être intégrée dans les fonctions d'archivage à chaque niveau du système d'archivage - nous tendons la main en interagissant avec d'autres, en apprenant à travailler ensemble, en menant des enquêtes rigoureuses sur les nouveaux développements, en transmettant les connaissances par la participation des élèves à la recherche. Les cours de formation continue sont importants, mais ils transmettent les connaissances déjà acquises. Nous avons besoin de programmes de formation archivistique plus solides qui intègrent des capacités de recherche et l'interdisciplinarité, et qui soutiennent la diffusion des connaissances des archives et l'acquisition de nouvelles connaissances lors de conférences organisées dans une variété d'industries. Nous n'apprenons pas ce qui est nécessaire pour relever les défis décrits plus haut en invitant un représentant de Google (ou Microsoft, ou Amazon) pour nous parler, mais plutôt en soumettant des propositions de communications pour des conférences parrainées des industries comme Google, par exemple le congrès annuel Global Cloud Security Alliance, le congrès annuel Digital Forensics, etc.

Comment l'information est-elle sauvegardée pour son utilité dans l'immédiat et à moyen terme à travers les technologies en considérant les importants changements qui se produisent? Comment les institutions de mémoire abordent-elles les questions posées par les nouvelles technologies en ce qui concerne leurs rôles traditionnels dans l'attribution de valeur, le respect des droits, et d'assurer l'authenticité et la fiabilité? Comment les institutions de mémoire peuvent-elles demeurer pertinentes comme source fiable d'information continue en profitant des possibilités de collaboration offertes par les nouveaux médias sociaux? Voir ACA Canadian Archival System Taskforce, Rapport intérimaire, juin 2013, p. 11.

La Stratégie nationale sur les infrastructures essentielles du Canada ⁴ (2009) identifie 10 secteurs d'infrastructures essentielles : l'énergie et les services publics, les finances, la nourriture, le transport, le gouvernement, les technologies de l'information et de communication, la santé, l'eau, la sécurité, la fabrication. Tous dépendent de documents authentiques, fiables, précis et qui sont conservés aussi longtemps que nécessaire. « Pour être efficace, la stratégie nationale doit être mise en œuvre en partenariat avec tous les niveaux de gouvernement et les secteurs des infrastructures essentielles ... Les gouvernements apportent de la valeur à la société à travers des activités telles que : donner aux propriétaires et aux exploitants de l'information opportune, exacte et utile sur les risques et les menaces ... » (souligné par l'auteur).

Un système d'archivistique canadien fort est un élément clé de cette infrastructure et devrait être proactif dans la détermination de sa place et de sa contribution à l'écosystème de l'infrastructure numérique du Canada. Le Leadership Council for Digital Infrastructure est certainement un forum dans le cadre duquel les archivistes doivent faire des propositions pour une infrastructure d'archivage intégrée.⁵ Pouvons-nous développer et soutenir une plateforme en nuage pour la communauté archivistique ? Les institutions et organisations créatrices de documents se tournent vers les plateformes en nuage en nombre croissant, car il peut simplifier les interactions avec les utilisateurs en réduisant le temps de traitement de l'information, en réduisant le coût des services et en améliorant la sécurité des données : il serait extrêmement difficile d'acquérir le contrôle de matériaux sélectionnés pour conservation permanente en les déplaçant dans un Dépôt numérique fiable (DNF), en particulier compte tenu de leurs quantités massives. En outre, il serait financièrement impossible pour chaque institution de disposer d'un DNF à l'interne (peut-être un seul mais qui serait limité dans ses fonctionnalités et dimensions), pour la manipulation de la conservation de données et documents qui doivent être strictement protégés en raison des lois sur la confidentialité et les droits juridiques, et ne peuvent pas être partagés même à travers des plateformes. Certes, avant de choisir l'option de plateforme en nuage, le système d'archivage devra répondre aux normes de service, de sécurité et de l'interopérabilité/réversibilité.⁶

⁴ Les infrastructures essentielles se réfèrent aux processus, systèmes, équipements, technologies, réseaux, biens et services essentiels à la santé, la sécurité ou le bien-être économique des Canadiens et au fonctionnement efficace du gouvernement. (NSCI 2009)

⁵ Voir <http://digitalleadership.ca/>.

⁶ Comme l'Australie (Archives nationales), la Nouvelle-Zélande (Bibliothèque nationale) et les États-Unis (organismes fédéraux), l'Europe se dirige clairement dans la direction de la plateforme en nuage : <http://www.futuregov.asia/articles/2013/nov/25/eu-cyber-security-agency-issues-cloud-deployment-g/>. Avec les pare-feu appropriés pour un contrôle séparé par chaque organisation responsable pour chaque ensemble de documents à un moment donné, le système archivistique canadien peut avoir une plateforme en nuage qui englobe tout plutôt qu'une plateforme en nuage gouvernemental distinct pour chaque palier de gouvernement ou même pour chaque organisme ou institution. Ce serait une mise en œuvre complète de l'idée totale des archives initialement exprimée par Wilfred I. Smith. Voir W.I. Smith, "Total Archives": The

Une plateforme en nuage archivistique canadien aurait besoin de fournir beaucoup plus de stockage, pour permettre toutes les fonctions d'archivage dans le cadre d'une répartition des responsabilités. En fait, en centralisant les fonds d'archives d'origine numérique au Canada dans le stockage en plateforme en nuage, nous pouvons éliminer la nécessité de concentrer les mêmes collections sous le contrôle de quelques grandes institutions. Les documents d'origine numériques des organismes et organisations qui ont un conservateur désigné (par exemple les organismes publics, les universités) seraient créés et conservés dans la plateforme en nuage des archives et resteraient sous le contrôle de leurs créateurs aussi longtemps qu'ils auraient besoin d'être utilisés, mais seraient surveillés par le conservateur désigné. Ensuite, le contrôle juridique et intellectuel des matériaux identifiés pour la conservation à long terme serait transféré au conservateur désigné.

Les créateurs de document, qu'il s'agisse d'organisations ou de personnes, qui n'ont pas de conservateur désigné, peuvent être dirigés vers des clauses types à inclure dans les contrats avec leurs fournisseurs, qui envisagent la possibilité de transfert de matériaux sélectionnés sur la plateforme en nuage de la communauté archivistique sous le contrôle des archives qui souhaitent les acquérir, et la faciliter. En fait, la collaboration avec l'industrie pour concevoir des modèles de contrats qui prennent en charge l'acquisition des archives a déjà commencé dans le cadre d'une variété de projets de recherche, et le Global Cloud Security Alliance surveille de développement de près.

Bien sûr, une grande réflexion est nécessaire sur la façon dont les fonctions archivistes seraient exercées, en particulier l'évaluation, compte tenu de la proportion gigantesque de la production éphémère. Cependant, il est déjà clair que, comme les créateurs des documents mettent davantage l'accent sur la communication que sur la mémoire, l'exécution sur l'enregistrement, les images sur les mots, et sur la réutilisation, la co-création, et le « crowdsourcing », la documentation du contexte devient vitale pour la compréhension et la préservation du sens, pour fournir une authentification collective des matériaux au sein de chaque fonds, et pour assurer un accès contrôlé et précis : la description archivistique reviendra au cœur de l'effort d'archives et deviendra de plus en plus complexe.

En effet, le système archivistique canadien est confronté à des défis qui semblent insurmontables, mais la résilience, la créativité et le dévouement de la communauté archivistique du Canada peut l'aider à relever le défi. C'est un moment très excitant pour une communauté aussi bien informée et engagée que celle des archivistes canadiens.

Canadian Experience," Dans Tom Nesmith, éd., *Canadian Archival Studies and the Rediscovery of Provenance* (Metuchen, 1993), p. 133.