

Données personnelles et confidentialité

Intervenant

Arnaud BELLEIL, expert en protection des données personnelles

Texte

Le *Big data* est un terme qui permet de réinterroger les questions entre l'archivage et les données personnelles. À première vue, on pourrait penser qu'il y a des points de vue opposés, voire contradictoires, entre ceux qui sont en charge de la mémoire et ceux qui doivent veiller à l'oubli. Finalement, il n'en est rien. La personne en charge de l'archivage, mais celle aussi qui a en charge le respect de la conformité par rapport à la loi Informatique et libertés, vont finalement avoir des pratiques qui vont dans le même sens. L'idée est de sélectionner ce qui doit être conservé et de le conserver pour une durée définie. Dans les deux cas, on va retrouver de bonnes pratiques d'archivage mais aussi des pratiques qui concourent à la protection de la vie privée, puisqu'on ne garde pas tout et que l'on garde pour une durée qui est délimitée.

Et ces deux pratiques vont un peu à l'opposé de nouvelles pratiques qu'on pourrait intégrer sous le vocable passe-partout de *Big data* qui est : « Gardons tout, gardons tout, gardons tout. On ne sait pas à quoi ça va servir mais ça servira bien un jour, et surtout ne jetons rien. »

Et là, on voit que données personnelles et archivage ne sont pas en opposition. Ils concourent plutôt à un même objectif par rapport à des aspects inquiétants de ce qu'on appelle le *Big data*.

La contradiction apparente entre mémoire numérique et droit à l'oubli peut réellement être considérée comme apparente et ça se confirme avec la doctrine de la CNIL (Commission nationale informatique et libertés) sur la conservation des archives électroniques. Il y a d'une part des textes sur les coffres forts numériques mais c'est plus des textes sur l'archivage que de grandes institutions mettent au bénéfice des particuliers, c'est un peu spécifique, et ce que dit la CNIL sur l'archivage électronique dans une recommandation de 2005, c'est que ce qui a trait à l'archivage doit être sorti du système de production, avoir un accès réservé à quelques personnes habilitées, que cet accès doit être tracé, éventuellement que les données puissent être anonymisées, tout ce qui concourt à faire de l'archivage non plus une menace mais finalement une forme d'outil de protection de la vie privée au regard des données informatiques.

Quand on associe les exigences d'archivage et de protection de la vie privée, on va soit se poser la question de la destruction des archives, soit aborder d'abord la question d'une autre façon. Une des façons d'aborder la question est d'assurer une confidentialité renforcée

des archives, notamment avec le chiffrement ou le cryptage des archives électroniques (une question qui prend de l'importance actuellement).

Traditionnellement, les archives ne sont pas chiffrées. Et ceci pour deux raisons :

- d'abord, cela porte atteinte à l'intégrité des documents numériques parce qu'on les transforme ;
- et l'autre question, qui est fondamentale, est que si on perd la clé de déchiffrement, on a perdu les archives, on a perdu l'accès à l'information, on n'a pas fait son travail d'archiviste.

Sous la pression de grands acteurs américains de l'économie numérique, des inquiétudes se font jour sur la confidentialité, par exemple sur les données de santé. Il y a donc une pression de la société pour plus de confidentialité qui va très vraisemblablement amener à chiffrer certaines archives, pas toutes mais certaines qui sont un peu critiques.

Certains techniques, qui commencent à se diffuser et qui peuvent paraître plus expérimentales, permettent de concilier l'exigence de conservation et la nécessité de protection des données personnelles et de confidentialité. Elles sont au nombre de trois :

1. **l'anonymisation** qui est un moyen de conserver l'information (sauf les données personnelles) mais il faut être cependant très attentif à ce qu'on appelle le risque de ré-identification, c'est-à-dire qu'en croisant des données potentiellement anonymes, on peut retrouver les personnes derrière ;
2. la deuxième technologie qui est parfois mise en œuvre, c'est distinguer l'archivage, qui est exhaustif, et la consultation avec une **diffusion qui peut être partiellement masquée** ;
3. et enfin la troisième solution, relativement innovante mais elle est déjà sur les réseaux sociaux et connaît un grand succès auprès des adolescents, et ça va peut-être se diffuser dans les organisations, c'est **l'information numérique à durée de vie limitée**. Un peu comme dans *Mission impossible*, au bout d'un certain temps ça disparaît tout seul, sans qu'on ait rien demandé et sans qu'on ait rien fait. C'est peut-être un rêve ou un cauchemar pour ceux qui seront en charge des archives numériques dans les prochaines années.