



de s'élargir, comme l'illustre l'arrivée des scanners et des imprimantes 3D.

Dès lors, on est tenté de penser que le problème ancien de la préservation des données est résolu... Après tout, quand il est possible de dupliquer, sans perte de qualité, et à coût quasi nul, autant de fois qu'on le souhaite, comment pourrait-on encore perdre de l'information ?

... n'existe pas !

Mais à y regarder de plus près, on s'aperçoit vite qu'il n'en est rien. Non seulement les techniques actuelles de reproduction n'ont pas résolu le problème, mais elles l'ont en fait compliqué. En premier lieu, les mêmes outils qui rendent facile la duplication de l'information ont aussi pour conséquence d'engendrer dans la pratique une quantité gigantesque d'information : pour l'année 2011, on a calculé que 1,8 zettaoctet de données ont été produites (voir la figure 2), et huit zettaoctets sont annoncés pour 2015 ! Précisons que la plus grande partie de cette information n'est pas créée directement par l'homme, mais par nos machines, qui gardent non seulement

les traces numériques de nos opérations informatiques, telles la transmission ou la conversion de fichiers, mais aussi les données et les métadonnées (données sur les données) produites par nos échanges sur les réseaux sociaux...

Décider de ce qui doit être préservé dans cette gigantesque masse et de ce qui peut tomber dans l'oubli n'est pas aisé. Ainsi, la Bibliothèque du Congrès des États-Unis archive désormais les messages envoyés par les citoyens américains sur le réseau *Twitter*, alors que ces commentaires et autres micro-informations devaient à l'origine n'être qu'éphémères...

Même si l'on disposait d'une méthode pour décider exactement de ce qui doit être préservé, la conservation à long terme des objets numériques serait bien plus problématique qu'il n'y paraît. À la différence des livres et des tableaux, qui livrent immédiatement leur contenu, l'accès aux objets numériques suppose un environnement technique complexe. Ce dernier comporte, d'une part, des supports matériels de stockage de l'information susceptibles de se dégrader et, d'autre part, des logiciels complexes, utilisant des formats définis et fonctionnant seulement sur certains processeurs qui, eux, sont susceptibles de devenir rapidement obsolètes. À moyen et long termes, les risques de pertes de données numériques sont donc considérables.

Illustrons cette réalité peu connue par un exemple : si l'on retrouvait aujourd'hui un document électronique rédigé il y a seulement... 30 ans, sans doute aurait-il la forme d'une disquette magnétique de 5 pouces 1/4 contenant un fichier *WordStar*. À supposer que cette grosse disquette ne soit pas trop dégradée, ce qui est fort peu probable (voir l'article de Franck Laloë et Erich Spitz dans ce numéro), nous nous heurterions à une autre difficulté : les lecteurs de tels supports ont presque tous disparu. De même, le logiciel *WordStar* n'est plus employé (car il n'est plus disponible), de sorte que le format dans lequel il enregistrait les données n'est plus reconnu par les traitements de texte actuels (voir la figure 4).

L'ESSENTIEL

■ La réplication numérique de données et leur diffusion par Internet semblent faciliter la conservation de la mémoire humaine.

■ En réalité, elle la menace, car les moyens techniques de cette réplication se périment et rendent les données inaccessibles.

■ Les offres d'hébergement gratuit des grands opérateurs d'Internet, pour utiles qu'elles soient, ne constituent pas non plus une solution sans risques.

■ Il reste à refaire sur le réseau ce que l'humanité a toujours fait : répartir sa mémoire entre tous !

© Shutterstock.com/festa

1. LE RÉSEAU HUMAIN est depuis toujours le dépositaire de l'expérience de l'humanité. Pour en faire aussi le dépositaire de sa mémoire numérique, encore faut-il faciliter les échanges de données entre pairs...