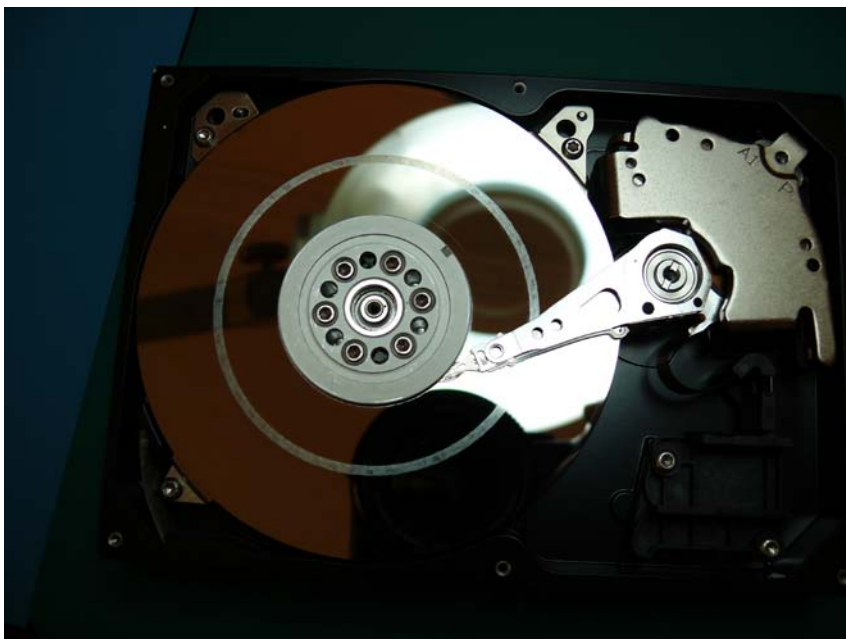




1. CE DISQUE DUR A ÉTÉ ENDOMMAGÉ par la tête de lecture qui a touché le matériau ferromagnétique à la surface du disque et tracé le sillon circulaire. Les données enregistrées sur le disque et situées dans le sillon sont définitivement perdues.

Qui possède encore un lecteur de disquette souple ou un lecteur de disque Zip ou Jaz ? De nombreux supports sont devenus obsolètes parce qu'ils ont été remplacés par des modèles plus performants ou parce qu'ils ont été des échecs commerciaux.

Ces deux aspects de la pérennité de l'information peuvent être réglés avec un peu d'organisation et de bon sens. Transposer au numérique le dépôt de métadonnées qui était déjà essentiel pour l'analogique n'est pas hors de portée ; de plus, on peut parfaitement choisir des systèmes d'exploitation, des logiciels et des formats de fichiers ouverts – dont le code est accessible et que chacun peut utiliser – ou très répandus, qui offrent des perspectives raisonnables de pérennité (voir l'article de Roberto Di Cosmo dans ce numéro).

Ces deux premiers problèmes sont importants, mais ils ne doivent pas masquer le troisième, qui est bien plus grave. On peut prendre toutes les précautions que l'on voudra en créant des métadonnées parfaites et en utilisant des formats robustes : si les données elles-mêmes disparaissent, cela ne sert évidemment à rien ! Ce problème a longtemps été sous-estimé et négligé. Hélas, nous allons voir que cet aspect de la pérennité des données pose des difficultés techniques et économiques qui sont loin d'être résolues.

La variété des dispositifs de stockage disponibles dans le commerce et leur évolution rapide peuvent donner l'impression qu'il existe de nombreuses méthodes différentes pour stocker et sauvegarder de l'information numérique. Certains supports sont-ils adéquats pour de l'archivage de longue durée ? En réalité, ces dispositifs fonctionnent sur un petit nombre de principes physiques et se rangent dans quelques catégories simples qui subissent toutes des détériorations en quelques années.

Tous les supports s'altèrent

La première catégorie fait appel à l'enregistrement magnétique : le système d'enregistrement est conçu pour orienter de façon contrôlée des petits domaines dans une couche magnétique déposée sur une surface plane. La lecture s'effectue ensuite par une tête électromagnétique sensible à cette orientation. C'est ainsi que fonctionnent les disques durs, les bandes magnétiques (ou LTO, pour *Linear Tape Open*) et même les disquettes de formats divers, dont la plupart ont maintenant disparu à cause du succès des mémoires flash. Les disques durs présentent de grandes capacités de stockage pour un coût faible, et avec un accès rapide aux données. Ils sont cependant fragiles et

soumis à des pannes soudaines et imprévisibles (ou *crash*) pouvant entraîner une perte totale des données (par exemple, si la tête de lecture touche la couche magnétique, voir la figure 1). Pour ne pas perdre ses données, une possibilité est la sauvegarde sur plusieurs disques en parallèle (tels les systèmes redondants RAID, pour *Redundant Array of Independent Disks*). La longévité des disques durs est difficile à évaluer, car elle dépend de nombreux facteurs indépendants. On s'accorde à penser néanmoins que, hors *crash*, une durée de vie de cinq ans est un bon ordre de grandeur. Pour optimiser cette longévité, il faut laisser tourner les disques durs en permanence, ce qui minimise les incidents. Les grands centres de données appliquent cette méthode coûteuse en électricité. On estime que les centres de données contribuent ainsi à plus de deux pour cent de la production de dioxyde de carbone d'origine humaine.

Il y a quelques années, les bandes magnétiques étaient considérées comme obsolètes, car moins performantes que les disques durs. Mais, précisément du fait de la nécessité de l'archivage des données, les bandes magnétiques ont effectué un retour spectaculaire sur la scène du numérique : elles conservent les données de façon passive, sans consommation électrique. Cependant, leur durée de vie est aussi limitée, de l'ordre de cinq ans, temps au bout duquel il devient indispensable d'effectuer une « migration » – c'est-à-dire de recopier les données sur une bande neuve, si on ne veut pas les perdre. Il existe d'ailleurs des robots dont la fonction est de gérer un grand nombre de bandes, d'y suivre le nombre d'erreurs et d'effectuer une migration vers des bandes neuves dès que c'est nécessaire.

La seconde catégorie de supports numériques fait appel à un enregistrement électrique et non magnétique : dans les mémoires flash, de très nombreux petits transistors stockent une charge électrique qui définit la valeur du bit enregistré. Les mémoires flash sont devenues d'usage universel, dans les appareils photo, les caméscopes, les ordinateurs, les téléphones, etc. Elles autorisent des accès aux données encore plus rapides que les disques durs, mais le prix par mégaoctet de données est plus élevé. Quant à la longévité des données stockées sur des mémoires flash, on dispose de peu d'informations à ce sujet. On sait cependant que les rayons cosmiques

déchargent progressivement les transistors et effacent donc les données. On considère souvent que la durée de vie est de l'ordre de cinq ans.

La troisième catégorie est constituée par les disques optiques numériques (DON), tels que les CD, les DVD et les disques *Blu-ray*. Les temps d'accès aux données enregistrées sont bien plus longs que ceux des disques durs et des mémoires flash. De plus, leur capacité est nettement inférieure à celle des disques durs. On fabrique pourtant chaque année des milliards de disques optiques numériques, encore que leur nombre soit en constante diminution chaque année; beaucoup pensent qu'ils sont en voie d'extinction progressive. En termes de longévité toutefois, ce sont les meilleurs supports disponibles actuellement : les disques « pressés » fabriqués en grande série ont une durée de vie qui serait comprise entre 20 et 30 ans. Cependant, ce n'est pas le cas des disques enregistrables, qui reposent sur des principes physiques plus délicats (inscription dans un colorant composé de molécules photosensibles ou, pour les disques réenregistrables, un changement de phase qui rend la surface plus ou moins réfléchissante). Les études du Laboratoire national de métrologie et d'essais en France et du Laboratoire lutherie acoustique musique de l'Université Paris 6 ont montré que certains disques enregistrés perdent leurs données en un an seulement ! D'autres, heureusement, ont des durées de vie dix fois plus longues.

Les disques optiques offrent des perspectives très intéressantes en termes de longévité. Il y a une vingtaine d'années une société française, *Digipress*, a commercialisé un CD ayant une excellente durée de vie, puisque les données étaient directement gravées dans du verre. Hélas, l'invention devait être trop en avance sur son temps, à une époque où beaucoup croyaient encore à la longévité des supports numériques. Ce produit n'a pas trouvé son marché. Mais, depuis deux ans, une société allemande, *Syplex*, a repris l'idée et propose des DVD gravés dans le verre dont la longévité est remarquable, comme l'ont montré des études du Laboratoire national de métrologie et d'essais (voir l'encadré page 82). L'avantage d'une telle solution est que les données sont stockées dans le format DVD, très répandu dans le monde, et qu'elle ne demande

aucun lecteur plus spécifique qu'un lecteur standard. Quelques autres produits sont proposés sur le marché et annoncés avec une grande longévité, mais à notre connaissance, aucun n'a subi des tests aussi soigneux que ces disques. Il faut espérer que les esprits aient maintenant suffisamment évolué pour que ce nouveau produit se répande rapidement, en particulier par des commandes institutionnelles, mais rien ne le garantit.

Graver dans le verre

Quelles stratégies peuvent être adoptées pour préserver l'information numérique à long terme avec les dispositifs existants ? Deux approches sont possibles : la stratégie dite passive, qui consiste à écrire l'information sur un support pérenne et à le stocker dans un endroit sûr ; la stratégie dite active, qui consiste à écrire l'information sur un support numérique standard et à en suivre constamment l'évolution, afin d'effectuer une migration vers un support neuf lorsque nécessaire. La stratégie passive est celle que l'humanité a pratiquée depuis toujours, mais elle utilisait des supports analogiques – que l'on songe à la bonne conservation des parchemins anciens, voire aux écritures cunéiformes sur des pierres ! Elle semble donc la plus naturelle, mais dans le cas du numérique, elle demande un investissement initial, par exemple pour la confection de DVD gravés dans le verre.

La stratégie active ne demande pas un tel investissement. Dans la plupart des cas,

LES AUTEURS

Franck LALÔÉ est directeur de recherche au Laboratoire Kastler Brossel de l'École normale supérieure. Il est président du Groupement d'intérêt scientifique SPADON. Il a été rapporteur du groupe Pérennité des supports numériques.

Erich SPITZ est membre de l'Académie des sciences et de l'Académie des technologies. Il a présidé le groupe Pérennité des supports numériques.

BIBLIOGRAPHIE

J. Zhang *et al.*, *5D data storage by ultrafast laser nanostructuring in glass*, compte rendu de conférence CLEO : Science and Innovations, 2013.

J. Perdureau, *Tenue dans le temps des disques enregistrables DVDR et DVD en verre (Syplex) à température et hygrométrie élevées*, rapport du GIS DON, LNE, 2012.

J.-C. Hourcade *et al.*, *Longévité de l'information numérique ; les données que nous voulons garder vont-elles s'effacer ?*, rapport à l'Académie des Sciences et à l'Académie des Technologies, EDP Sciences, 2010.

L'altération, un processus mal compris

Les différents supports numériques ont des durées de vie relativement courtes. De nombreux facteurs entrent en jeu et il est difficile de dire précisément quels phénomènes sont en cause et lesquels sont les plus importants. Pour les disques durs, les pannes mécaniques sont une difficulté majeure. Mais d'autres problèmes peuvent surgir : certains éléments du disque stockant l'information peuvent devenir défectueux et illisibles.

Les rayons cosmiques peuvent décharger les transistors des mémoires flash et modifier l'infor-

mation. Les supports optiques, les CD et les DVD, correspondent à un empilement de couches de protection en plastique, de couches contenant les données et d'une pellicule métallique qui sert de réflecteur pour le laser de lecture des données. La pellicule plastique est sensible à la surexposition aux ultraviolets de la lumière du Soleil et à la chaleur. Les tests de longévité de ces supports ont aussi montré qu'ils sont sensibles aux réactions d'oxydation par l'eau et par l'air. Les supports optiques enregistrables sont encore moins stables.

Les facteurs externes ne sont pas les seuls à influencer la durée de vie des supports. La qualité des matériaux utilisés et les précautions prises lors de la fabrication sont aussi importantes. Le Laboratoire national de métrologie et d'essais a montré que sept pour cent des DVD enregistrables du commerce présentaient des microtaches. Elles seraient dues à des impuretés présentes dans la couche réfléchissante et accéléreraient le processus d'oxydation, réduisant la durée de vie du support à un an.