

Recommandations du GIS-DON concernant la préservation à long terme des données numériques (textes, documents, photos, enregistrements sonores, vidéos, etc.).

Trois étapes sont à prendre en compte ; les deux premières sont relativement simples, mais néanmoins indispensables. La troisième est plus délicate, dans la mesure où elle demande un suivi constant ; cependant c'est la clé du problème et si elle est ignorée, les autres précautions s'avéreront inutiles à terme.

1 La première étape découle du simple bon sens :

(a) il faut savoir trier l'information destinée à une préservation à long terme, en choisissant le plus essentiel. Il n'est pas nécessairement utile de préserver toute l'information brute qui nous passe sous les yeux, dont beaucoup n'a qu'un intérêt éphémère.

(b) un document n'est utile et intéressant à long terme que s'il est accompagné d'informations décrivant son contenu. Par exemple, les photos anciennes ne gardent tout leur intérêt que si l'on sait de qui il s'agit (par exemple position d'un ancêtre dans l'arbre généalogique, lieu de prise de vue, etc.). De façon générale, la description des contenus est essentielle.

Il est recommandé d'organiser les documents en répertoires qui en fournissent un classement naturel et clair, top d'une consultation des décennies plus tard: pour des photos par exemple, on les rangera dans des répertoires correspondant aux dates, ou aux lieux, etc.

(c) on évitera évidemment les formats de fichiers qui évoluent vite ou sont associés à des logiciels changeant constamment ; les formats ouverts et très répandus seront préférés ; pour l'audio et la vidéo, on évitera autant que possible la compression et les formats dont les paramètres de réglage et les standards sont nombreux, variables, et souvent assez rapidement obsolètes.

Pour le format PDF par exemple (format ouvert, mais pas libre), on privilégiera PDF-A, version épurée d'un certain nombre de fonctionnalités peu compatibles avec l'archivage à long terme. De même, pour conserver une " présentation Powerpoint ", on renoncera aux animations et aux effets spéciaux, pour la transformer en un pdf qui soit plus sûr et moins dépendant de la version du logiciel. Chaque fois que possible, dans les domaines scientifiques on utilisera le format Tex/Latex, qui est probablement un de ceux dont la durabilité attendue est la plus grande.

2. Il convient de se prémunir contre les accidents soudains : un disque dur cesse brusquement de fonctionner (crash) une clé USB est perdue ou volée, un incendie ou une inondation survient, etc. La plupart de ces risques sont d'ailleurs communs à l'analogique et au numérique. Il convient donc de faire plusieurs copies (deux, voire trois) placées en des endroits différents, sur des supports différents.

Si tous sont des disques durs, on privilégiera des marques différentes, mais il est encore préférable de varier les types de supports. Pour les CD-R et les DVD-R, les tests montrent qu'il est difficile de recommander telle ou telle marque, la qualité de la fabrication variant dans le temps pour chacune d'entre elles. En particulier, les mesures montrent que les supports dits " archivage " ne sont pas nécessairement meilleurs que les autres.

3. Aucun support d'information numérique ne peut actuellement garantir la conservation de l'information au delà de quelques années. Même en l'absence de toute lecture, ils se dégradent inexorablement dans le temps. En attendant l'apparition de supports dont la durée de vie atteigne plusieurs décennies, voire des siècles, la seule solution pour le moment reste la recopie régulière, associée si possible à un contrôle régulier de la qualité des données. Heureusement l'utilisation

systematique et automatique des codes de corrections d'erreurs permet des recopies parfaites, tant que le support numérique est en bon état ; en revanche, à partir d'un certain taux de dégradation, on bascule brusquement d'une recopie parfaite vers une recopie impossible.

Ainsi il est recommandé de recopier à l'identique tous les 5 ans ou mieux encore tous les 3 ans pour plus de sûreté, le contenu d'un support ancien sur un support neuf. Attendre 10 ans comporte un risque réel, quel que soit le support Si l'on a retenu la solution de trois copies, on pourra par exemple en recopier tour à tour une des trois chaque année.

Si le support utilisé est un disque optique enregistrable, ne pas graver la fin du disque (les derniers 10%), plus sensible au vieillissement que le reste du disque

Conclusion : si vous appliquez cette méthode de façon systématique, profitant de ce que le monde du numérique permet des recopies sans erreur, tout en étant demandant relativement peu de place de stockage, vous êtes pratiquement assurés de transmettre vos documents aux générations futures. Mais, pour le moment, et dans l'état actuel des produits disponibles, avec un peu de malchance négliger une seule des étapes peut suffire à déclencher une perte totale et irréversible. L'information numérique dont personne ne s'occupe, stockée pendant des années au fond d'une armoire, n'a que bien peu de chances d'être encore lisible lorsqu'on voudra y avoir accès.