

**DIRECTIVES POUR LA SAUVEGARDE  
DU PATRIMOINE NUMERIQUE**



**Document élaboré par la Bibliothèque nationale d'Australie**

Division de la société de l'information

Organisation des Nations Unies pour l'éducation, la science et la culture

---

## TABLE DES MATIERES

Page

Remerciements.....	1
Préface .....	3

### 1re PARTIE. ELEMENTS D'INTRODUCTION

Chapitre 1	Introduction .....	7
Chapitre 2	Le Projet de Charte de l'UNESCO sur la préservation du patrimoine numérique .....	9
Chapitre 3	Guide des directives.....	15
Chapitre 4	Note sur la terminologie .....	19
Chapitre 5	Résumé des principes .....	21

### 2e PARTIE. PERSPECTIVES DE GESTION

Chapitre 6	Comprendre le patrimoine digital.....	27
Chapitre 7	Comprendre la préservation numérique .....	33
Chapitre 8	Comprendre les programmes de préservation numérique .....	37
Chapitre 9	Prendre des responsabilités.....	43
Chapitre 10	Gérer les programmes de préservation numérique .....	53
Chapitre 11	Travailler ensemble .....	65

### 3e PARTIE. PERSPECTIVES TECHNIQUES ET PRATIQUES

Chapitre 12	Décider de ce que l'on conserve .....	73
Chapitre 13	Travailler avec les producteurs.....	81
Chapitre 14	Prendre le contrôle : transfert et métadonnées.....	93
Chapitre 15	Gérer les droits.....	109
Chapitre 16	Protéger les données .....	117
Chapitre 17	Maintenir l'accessibilité .....	131
Chapitre 18	Quelques points de départ.....	161

### 4e PARTIE. INFORMATIONS COMPLEMENTAIRES

19	Glossaire .....	171
20	Bibliographie .....	174
21	Index .....	193

Site Web : [www.unesco.org/webworld/mdm](http://www.unesco.org/webworld/mdm)

Contact : [a.abid@unesco.org](mailto:a.abid@unesco.org)

## Remerciements

Les Directives pour la préservation du patrimoine numérique ont été principalement élaborées par Colin Webb, avec l'aide et la contribution de plusieurs autres membres du personnel de la Bibliothèque nationale d'Australie, dont Kevin Bradley, Debbie Campbell, Gerard Clifton, Mark Corbould, Maura O'Connor, Margaret Phillips et Julie Whiting, qui ont apporté des idées et dans certains cas rédigé un projet écrit pour quelques chapitres. Elles ont bénéficié aussi du concours d'un certain nombre de personnes extérieures à la Bibliothèque nationale d'Australie qui ont apporté idées et commentaires, dont le professeur Arnaldo Coro Antich de La Havane, le docteur T. Matthew Ciolek de l'Université nationale d'Australie, M. Simon Davis des Archives nationales d'Australie, M. Ian Gilmour de Screen Sound, Australie, le docteur Henry Gladney de Californie, M. Roger Harris de Hong Kong, Mme Justine Heazlewood du VERS, le docteur Graeme Johanson de l'Université Monash, Mme Maggie Jones de la Digital Preservation Coalition (Royaume-Uni), Mme Anne Kenney de l'Université Cornell ; M. Stephen Knight de la Bibliothèque nationale de Nouvelle-Zélande ; le docteur Simon Pockley du Centre australien de l'image animée, le docteur Johan Steenbakkers de la Koninklijke Bibliotheek, M. Hiroyuki Taya de la Bibliothèque de la diète nationale, Japon, M. Paul Trezise de GeoScience, Australie, et Mme Deborah Woodyard de la British Library.

Je tiens aussi à saluer l'attention et les apports que j'ai reçus des participants aux diverses réunions de consultation régionale, ainsi que les conseils éclairés et la patience de M. Abdelaziz Abid de la Division de la société de l'information, UNESCO, Paris.

Si je n'ai pas toujours été en mesure de reproduire leurs commentaires dans les directives, j'ai beaucoup appris en travaillant avec chacun d'entre eux.

Une grande partie du contenu de ces Directives se fonde aussi sur le travail d'autres personnes fort averties qui travaillent sur les programmes de préservation et de recherches dans le monde entier : il aurait été difficile de se passer de leur aide. Pour ne pas encombrer le texte avec des citations, les noms et les sources apparaissent généralement dans la liste bibliographique, sauf lorsqu'il y a un lien direct et exclusif entre un commentaire des Directives et une source spécifique. Il va de soi que la contribution qu'elles ont apportées à leur insu à ces Directives est d'une importance considérable et que je leur en sais gré.

A tous ceux qui m'ont apporté leur contribution sous une forme ou une autre, j'exprime ma profonde reconnaissance. Les mauvaises interprétations, les fausses perspectives, les ambiguïtés ou les erreurs sont presque à coup sûr de mon seul fait.

Colin Webb  
Directeur de la préservation  
Bibliothèque nationale d'Australie  
Mars 2003

## Préface

Une large part des vastes quantités d'informations produites dans le monde est d'origine numérique et se présente sous une grande diversité de formats : texte, base de données, son, film, image. Pour les institutions culturelles qui ont traditionnellement la charge de recueillir et de préserver le patrimoine culturel, il est devenu extrêmement urgent de savoir quels sont les matériaux qu'il convient de conserver pour les générations futures et comment il faut procéder à leur sélection et à leur préservation. Ce gigantesque trésor d'informations numériques produit aujourd'hui dans presque tous les domaines de l'activité humaine et conçu pour être accessible sur ordinateur risque d'être perdu si l'on ne développe pas des techniques et des politiques spécifiques en vue de le conserver.

Préserver des informations scientifiques précieuses, des données de recherche, des productions des médias, de l'art numérique, pour ne citer que quelques champs d'activité, soulève de nouvelles difficultés. Si l'on veut avoir accès à ce matériau sous sa forme originale, il faut maintenir l'équipement technique - matériel et logiciel originaux ou compatibles - en même temps que les fichiers numériques qui constituent les données concernées. Souvent les composants multimédia des sites Web, notamment les liaisons Internet, présentent une complication supplémentaire pour ce qui est du copyright et de la géographie : il devient parfois difficile de déterminer à quel pays appartient un site Web.

L'UNESCO étudie ces questions en vue de définir une norme pour guider les efforts de préservation qu'entreprennent les gouvernements à l'âge numérique. La Conférence générale a adopté la résolution 34 lors de sa 31<sup>e</sup> session, attirant ainsi l'attention sur l'augmentation croissante du patrimoine numérique dans le monde et sur la nécessité de lancer une campagne internationale pour sauvegarder la mémoire numérique des dangers qui la menacent. La Conférence générale avait aussi invité le Directeur général à préparer, pour la session du printemps 2001 du Conseil exécutif, un document à débattre qui comporte une ébauche visant à la fois à préserver des documents d'origine numérique et à encourager les organisations gouvernementales et non gouvernementales ainsi que les institutions internationales, nationales et privées, à garantir que les politiques nationales donnent une priorité absolue à la préservation du patrimoine numérique.

Au cours de la réunion du Conseil exécutif de l'Organisation en mai 2001, les Etats membres se sont mis d'accord sur la nécessité d'une action rapide pour sauvegarder le patrimoine numérique. Le débat s'est largement inspiré d'un document de discussion établi pour l'UNESCO par la European Commission on Preservation and Access (ECPA), une fondation à but non lucratif basée à Amsterdam, qui a esquissé les problèmes inhérents à la préservation numérique.

Les méthodes de préservation traditionnelles, tel le "dépôt légal" en vigueur dans les bibliothèques nationales pour garantir la conservation d'un exemplaire de tous les matériaux imprimés, ne sont pas applicables comme telles au matériau numérique pour diverses raisons, notamment parce que les "publications" du Web utilisent des données mémorisées sur des serveurs dans différentes parties du monde. Le simple volume des données concernées fait également problème. L'Internet, estime-t-on, compte un milliard de pages d'une durée de vie moyenne extrêmement brève (évaluée entre 44 jours et deux ans).

L'Internet, considéré comme le support de publication le plus démocratique qui ait jamais existé, est en extension croissante et mérite, selon certains, d'être préservé intégralement si l'on considère ses pages et ses forums de discussion comme un miroir inestimable de la société.

Garantir que le matériel numérique sauvegardé dans les archives reste accessible sous sa forme originale pose des problèmes techniques. La part de l'information et de l'art produits à travers le monde sur des supports traditionnels - page imprimée, bande analogique ou film diminue chaque année si on la compare au matériel conçu pour être accessible sur ordinateur : logiciels et matériels sont constamment remplacés par des générations nouvelles plus puissantes qui ne tardent pas à devenir incompatibles avec leurs prédécesseurs. Résultat : au bout de quelques années seulement, ce matériel - qui comprend souvent du son et des graphiques ou des images animées, ainsi que des liens avec des sites de l'Internet et/ou des bases de données - devient inaccessible.

Le simple volume de données à passer au crible pour sélectionner ce qui mérite d'être préservé est confondant. "Le montant annuel de la production mondiale de matériaux imprimés, filmiques, optiques et magnétiques nécessiterait environ 1,5 milliard de gigaoctets d'archivage. C'est l'équivalent de 250 mégaoctets par personne pour chaque homme, femme et enfant sur la terre," selon une étude récente réalisée par l'Ecole de gestion de l'information et des systèmes à l'Université de Californie de Berkeley (2).

Un autre problème complexe touche au copyright, y compris le copyright du logiciel nécessaire pour accéder aux fichiers numériques. Des sites Web qui offrent un mélange de matériaux de sources diverses peuvent ainsi présenter un ensemble ahurissant de droits. L'accord sur le principe du "droit de reproduire pour sauvegarder" reste encore à mettre au point à l'échelle mondiale.

Alors que de nombreux pays ont pris des initiatives précieuses pour préserver le patrimoine numérique, y compris les sites Web, l'étude de l'ECIPA montre les limites de ces efforts et plaide en faveur de normes internationales.

Pour résoudre la complexité des problèmes en jeu, la tâche de préservation exige d'impliquer les fabricants de l'information numérique, logiciels compris : ils devraient prendre en compte la conservation lorsqu'ils conçoivent leurs produits. Le temps n'est plus où la préservation était du ressort exclusif des institutions d'archivage.

Coopération, orientation, direction et partage des tâches sont des éléments clés pour la sauvegarde du patrimoine numérique. Les institutions culturelles ont besoin de la coopération des créateurs de l'information et des fabricants de logiciels. Des ressources et un soutien adéquats au plan politique sont indispensables pour permettre aux générations futures de continuer à avoir accès à la richesse des ressources numériques dans la création desquelles nous avons tant investi au cours des dernières décennies.

L'UNESCO a élaboré une stratégie fondée sur les constatations précédentes pour encourager la préservation numérique. Cette stratégie a pour axes principaux : (a) une vaste opération de consultation avec les gouvernements, les acteurs politiques, les producteurs de l'information, les institutions et les experts du patrimoine, l'industrie des logiciels ainsi que les organisations qui fixent les normes ; (b) la diffusion de directives techniques ; (c) la mise en place de projets pilotes ; (d) et la préparation d'un projet de charte sur la préservation du patrimoine numérique qui puisse être adopté par la Conférence générale lors de sa 32e session.

Le présent document, élaboré pour l'UNESCO en contrat avec la Bibliothèque nationale d'Australie, présente des directives générales et techniques pour la préservation et l'accès permanent au patrimoine numérique mondial en constante expansion. Il est destiné à accompagner le projet de charte sur la préservation du patrimoine numérique.

Je remercie Colin Webb et la Bibliothèque nationale d'Australie d'avoir élaboré ces directives et organisé la réunion de Consultation régionale sur la préservation du patrimoine numérique pour l'Asie et le Pacifique, qui s'est tenue à Canberra (Australie) du 4 au 6 novembre 2002. Ce fut la première d'une série de réunions de consultations régionales similaires qui ont eu lieu à Managua (Nicaragua) du 18 au 20 novembre 2002, à Addis-Abeba (Ethiopie) du 9 au 11 décembre 2002, à Riga (Lettonie) du 18 au 20 décembre 2002, et à Budapest (Hongrie) les 17 et 18 mars 2003.

A ces réunions régionales ont participé quelque 175 experts originaires de 86 pays, représentant une large gamme de personnes et de disciplines concernées, notamment des bibliothécaires et des archivistes, des prestataires de service de l'Internet, des agences nationales de normalisation, des représentants de l'industrie du logiciel et du matériel informatique, des journalistes, des juristes, des autorités universitaires et gouvernementales. Ils ont tous enrichi de leurs commentaires et réflexions le projet des directives et le projet de charte préliminaire sur la préservation du patrimoine numérique.

Nous espérons que ces Directives se révéleront utiles en aidant les responsables et les spécialistes de la préservation à aborder la complexité des problèmes techniques complexes propres à la préservation et l'accessibilité permanente du patrimoine numérique mondial.

Abdelaziz Abid  
Division de la société de l'information  
UNESCO

- 1 <http://unesdoc.unesco.org/images/0012/001255/125523e.pdf>
- 2 <http://www.sims.berkeley.edu/how-much-info>

**PREMIERE PARTIE**  
**ELEMENTS D'INTRODUCTION**

## Chapitre 1. Introduction

Notre patrimoine culturel, scientifique et informatif existe de plus en plus sous des formes numériques et cela de plus en plus exclusivement. Les technologies utilisées pour créer et profiter du patrimoine numérique présentent de nombreux avantages, d'où leur adoption extrêmement rapide dans de nombreuses parties du monde.

Mais il y a de très sérieux défis à relever si l'on veut garder notre patrimoine numérique naissant, mais déjà proliférant, utilisable et disponible. Les supports que nous utilisons pour le transporter et le stocker sont instables, et la technologie nécessaire pour y accéder rapidement ne cesse d'être remplacée, vague après vague, par des technologies nouvelles. L'usage de ces technologies se perdant, l'accès au patrimoine numérique qui en était tributaire est également perdu.

Ces défis, loin d'être seulement de nature technique, touchent à des questions d'organisation et de société puisque nous sommes aux prises avec la responsabilité de garder les lignes d'accès ouvertes sur de longues périodes de temps, souvent avec des ressources insuffisantes et des stratégies incertaines.

L'intérêt que porte l'UNESCO à cette situation n'a rien de surprenant. Cette Organisation internationale existe en partie pour encourager et permettre la préservation et la jouissance du patrimoine culturel, scientifique et informatif des peuples du monde entier. Elle ne pouvait négliger l'augmentation du patrimoine numérique et sa vulnérabilité.

Ces Directives ne constituent qu'une part de la campagne à long terme menée par l'UNESCO pour améliorer l'accès au patrimoine numérique chez tous les peuples du globe, et pour garantir que chaque communauté a entre les mains les moyens de préserver son patrimoine numérique.

La portée et l'ambition des Directives sont limitées. Dans un domaine où l'évolution est si rapide, mais dont le champ est déjà vaste et complexe, elles ne peuvent présenter qu'une petite quantité d'informations. Les buts recherchés sont de fournir des indications à des individus et des organisations qui envisagent d'assumer leurs responsabilités dans la préservation du patrimoine numérique, ceux-ci étant souvent placés dans une situation caractérisée à la fois par une absence de moyens et une pléthore d'informations. Aussi a-t-on délibérément adopté une façon d'aborder les principes qui se présente comme un récapitulatif (assez étendu) des problèmes et des possibilités à prendre en considération dans les programmes.

Il est impossible de répondre à tous les problèmes techniques et pratiques qui surgiront dans la gestion des programmes de préservation numérique. Il sera donc plus utile de voir dans ces Directives un guide des questions auxquelles les responsables de programme ont besoin de trouver des réponses. Mais ce guide s'appuie sur la ferme conviction qu'il est temps de poser les questions qui débouchent sur une action positive, plutôt que de continuer à poser celles qui ne font que souligner les difficultés.

Il faut espérer que ces directives, en complément avec la richesse des informations techniques disponibles dans les sources fournies par la liste bibliographique, aideront les responsables de programme de préservation à clarifier les décisions et les mesures à prendre, les principes à retenir et les considérations pratiques à envisager.

Le public auquel ces Directives s'adressent comprend des organisations culturelles et de recherche comme les bibliothèques, les archives, les musées, les instituts de recherche, les centres d'archivage de données, les éditeurs, les groupes communautaires et autres personnes qui peuvent être amenées à prendre des responsabilités dans la préservation du patrimoine numérique et qui y trouveront leur intérêt. Il inclut de nombreuses personnes qui rassemblent et préservent depuis longtemps le patrimoine de documents, d'articles, de publications, de cartes, de manuscrits, d'illustrations, d'enregistrements sonores, d'imagerie animée, d'objets culturels et d'informations scientifiques, statistiques et de recherche de la "mémoire" du monde. Il comprendra aussi un grand nombre d'autres personnes qui viennent à la préservation numérique dans un contexte différent et qui sont moins familiers avec les perspectives de préservation élaborées dans les organisations de la "mémoire".

Ces Directives ont été élaborées par la Bibliothèque nationale d'Australie en contrat avec l'UNESCO et se fondent sur une vaste documentation, l'expérience propre de la Bibliothèque et les consultations organisées par l'UNESCO dans différents centres régionaux. Pour plus d'information sur les contributions et les titres des personnes qui ont apporté leur collaboration, le lecteur consultera la page des remerciements ; pour être aidé dans l'utilisation de ces directives, il se reportera au chapitre 3 : Guide des directives.

## **Chapitre 2. Le Projet de Charte de l'UNESCO sur la préservation du patrimoine numérique**

### INTRODUCTION

Le projet de charte de l'UNESCO sur la préservation du patrimoine numérique présente un modèle contraignant de préservation numérique. Il est inclus dans les Directives pour établir un lien clair entre les deux documents et pour traiter de sujets touchant au soutien de principe et aux politiques publiques qui n'entrent pas dans le champ des directives techniques et pratiques.

### PROJET DE CHARTE REVISE SUR LA PRESERVATION DU PATRIMOINE NUMERIQUE

#### **PREAMBULE**

La Conférence générale,

*Considérant* que la disparition du patrimoine sous quelque forme que ce soit constitue un appauvrissement du patrimoine de toutes les nations,

*Rappelant* que la Constitution de l'UNESCO stipule que l'Organisation maintiendra, augmentera et divulguera la connaissance, en assurant la conservation et la protection du patrimoine mondial de livres, oeuvres d'art et monuments de l'histoire et de la science, que son Programme d'"Information pour tous" fournit une plate-forme de discussions et d'action pour les politiques d'information et de sauvegarde de la connaissance enregistrée, et que son programme "Mémoire du monde" vise à assurer la préservation et l'accessibilité universelle du patrimoine documentaire mondial,

*Reconnaissant* que ces ressources d'information et d'expression créatrice sont de plus en plus produites, diffusées, consultées et conservées sous forme numérique, créant dès lors un nouvel héritage, le patrimoine numérique,

*Conscients* du fait que l'accès permanent à ce patrimoine offrira des possibilités accrues pour la création, la communication et le partage des connaissances parmi tous les peuples, ainsi que pour la protection des droits et le soutien de la responsabilité,

*Comprenant* que ce patrimoine numérique risque d'être perdu et que sa préservation au profit des générations actuelles et futures pose un problème urgent d'un intérêt universel,

*Gardant à l'esprit* la Déclaration universelle sur la diversité culturelle de l'UNESCO,

*Proclame* les principes suivants et *adopte* la présente Charte.

## LE PATRIMOINE NUMERIQUE COMME PATRIMOINE COMMUN

### **Article premier - Le patrimoine numérique**

Les ressources des connaissances ou de l'expression humaines, qu'elles soient culturelles, éducatives, scientifiques et administratives, ou qu'elles englobent des informations techniques, juridiques, médicales et autres, sont de plus en plus créées numériquement, ou converties sous forme numérique à partir de ressources analogiques existantes. Là où les ressources sont d'origine numérique, il n'existe pas d'autre format que l'original numérique.

Les matériaux numériques comprennent des textes, des bases de données, des images fixes et animées, du son, des graphiques, des logiciels, des pages Web dans une gamme étendue et croissante de formats. Ils sont souvent éphémères, et leur conservation exige une fabrication, un entretien et une gestion avisés.

Nombre de ces ressources ont une valeur et une importance durables, et constituent donc un patrimoine qui doit être protégé et préservé pour les générations actuelles et futures. Ce patrimoine peut exister dans n'importe quelle langue, n'importe quelle partie du monde et n'importe quel domaine des connaissances ou de l'expression humaines.

### **Article 2 - L'accès au patrimoine numérique**

L'objectif de la préservation du patrimoine numérique est de garantir la permanence de son accès. Ainsi, l'accès aux matériaux du patrimoine numérique, en particulier ceux du domaine public, doit être équitable et exempt de toutes restrictions non fondées. En même temps, la sécurité des informations confidentielles et privées doit être protégée contre toute forme d'intrusion.

Il revient à chaque Etat membre de coopérer avec les organisations et les institutions compétentes pour encourager un environnement juridique et pratique qui accroîtrait au maximum l'accessibilité au patrimoine numérique. Un équilibre juste entre les droits légitimes des créateurs et autres détenteurs de droits et ceux du public à l'accès aux matériaux du patrimoine numérique doit être réaffirmé et encouragé.

## SE PROTEGER CONTRE LA PERTE DU PATRIMOINE

### **Article 3 - La menace de la perte**

Le patrimoine numérique mondial court le risque d'être perdu pour la postérité. Les facteurs responsables comprennent l'obsolescence rapide du matériel et du logiciel informatiques qui lui donne naissance, les incertitudes qui pèsent sur les ressources, les responsabilités et les méthodes concernant l'entretien et la préservation, et l'absence d'une législation de soutien.

Le changement d'attitude a pris du retard sur le changement technologique. L'évolution numérique a été trop rapide et trop coûteuse pour les gouvernements et les institutions qui n'ont pu, de ce fait, élaborer des stratégies de préservation en temps voulu et à jour. On n'a pas encore pris pleinement mesure des menaces qui pèsent sur le potentiel économique, social, intellectuel et culturel de ce patrimoine - poutre maîtresse du futur.

#### **Article 4 - La nécessité d'agir**

A moins que ne soient contrées les menaces principales, la perte du patrimoine numérique sera rapide et inévitable. Il est urgent d'éveiller les consciences et de défendre cette cause, en alertant les acteurs politiques et en sensibilisant le grand public à la fois au potentiel du support numérique et aux aspects pratiques de la préservation. Les Etats membres gagneront à encourager des mesures juridiques, économiques et techniques pour sauvegarder le patrimoine.

#### **Article 5 - La continuité de l'information numérique**

Le patrimoine numérique fait partie du continuum plus large de l'information numérique. Pour préserver le patrimoine numérique, des mesures doivent être prises tout au long de la durée de vie de l'information. La préservation du patrimoine numérique commence avec la conception de systèmes fiables qui produiront des objets numériques authentiques et stables.

### **LES MESURES NECESSAIRES**

#### **Article 6 - Elaborer des stratégies et des politiques**

On peut élaborer des stratégies et des politiques pour préserver le patrimoine numérique en tenant compte du degré d'urgence, des circonstances locales, des moyens disponibles et des projections dans le futur. Cette tâche sera facilitée par la coopération des créateurs, des détenteurs de copyright et de droits connexes, et des institutions habilitées à établir des normes et des compatibilités communes, et par le partage des ressources.

#### **Article 7 - Définir ce qui doit être gardé**

Comme pour tout patrimoine documentaire, les principes de sélection peuvent varier selon les pays, même si les principaux critères pour guider le choix des matériaux numériques à garder doivent être leur importance et la valeur durable qu'ils présentent sur le plan culturel, scientifique, testimonial ou autre. Les décisions de sélection et toutes révisions ultérieures doivent être mises en oeuvre d'une manière responsable et justifiable et fondées sur des politiques, des procédures, des normes et des principes définis.

#### **Article 8 - Protéger le patrimoine numérique**

Les Etats membres ont besoin de cadres appropriés pour assurer la protection de leur patrimoine numérique. Les forces de marché n'y parviendront pas seules.

Elément clé de la politique de préservation nationale, la législation des archives et le dépôt légal ou spontané dans les bibliothèques, les archives, les musées et autres gisements de données doivent inclure le patrimoine numérique. La législation du copyright et des droits connexes doit permettre à de telles institutions d'entreprendre légalement des procédures de préservation.

Le droit à l'accès permanent des matériaux du patrimoine numérique légalement déposés, dans des limites raisonnables, doit être garanti sans porter préjudice à leur exploitation normale.

Des cadres juridiques et pratiques pour garantir son authenticité sont cruciaux pour éviter une manipulation ou une modification intentionnelle du patrimoine numérique. Ils réclament que le contenu, la fonctionnalité des dossiers et de la documentation soient entretenus à un niveau suffisant pour mettre en sécurité une pièce d'archive authentique.

## **Article 9 - Encourager la diversité culturelle**

Le patrimoine numérique est, de par sa nature, non limité par le temps, la géographie, la culture et le format. Il est propre à une culture, mais potentiellement accessible à toute personne dans le monde. Les minorités peuvent s'adresser aux majorités, les individus à un public universel.

Le patrimoine numérique de toutes les régions, tous les pays et communautés doit être préservé et rendu accessible, en créant avec le temps une représentation équilibrée et équitable de tous les peuples, toutes les nations, les cultures et les langues.

## **LES RESPONSABILITES**

### **Article 10 - Les rôles et les responsabilités**

Il incombe à chaque Etat membre de désigner un ou plusieurs organismes pour assumer la responsabilité de la coordination de la préservation du patrimoine numérique, et de fournir le personnel et les ressources nécessaires. Le partage des tâches et des responsabilités peut être fondé sur des rôles et des compétences existants.

Il convient de prendre des mesures pour :

- (a) exhorter les programmeurs de matériel et de logiciels informatiques, les éditeurs, les producteurs et les distributeurs de matériaux numériques ainsi que d'autres partenaires du secteur privé à coopérer avec les bibliothèques nationales, les archives, les musées et autres organisations publiques du patrimoine pour préserver le patrimoine numérique ;
- (b) développer la formation et la recherche, et faire partager l'expérience et les connaissances entre les institutions et les associations professionnelles concernées ;
- (c) encourager les universités et autres organismes de recherche à assurer la préservation des données de recherche.

### **Article 11 - Partenariats et coopération**

La préservation du patrimoine numérique réclame des efforts soutenus de la part des gouvernements, des créateurs, des éditeurs, des industries concernées et des institutions du patrimoine.

Face au fractionnement numérique qui règne actuellement, il est nécessaire de renforcer la coopération et la solidarité internationales pour permettre à tous les pays d'assurer la création, la diffusion, la préservation et l'accessibilité permanente de leur patrimoine numérique.

Les industries, les éditeurs et les médias de communication de masse sont incités à encourager et partager les connaissances et les compétences techniques.

La stimulation suscitée par des programmes d'éducation et de formation, les dispositions prises pour un partage des ressources, et la diffusion de résultats de recherche et de pratiques plus adéquates démocratiseront l'accès aux techniques de la préservation numérique.

## **Article 12 - Le rôle de l'UNESCO**

L'UNESCO, en vertu de son mandat et de ses fonctions, a la responsabilité de :

- (a) prendre en compte les principes énoncés dans cette Charte pour le fonctionnement de ses programmes et d'encourager leur mise en oeuvre au sein du système des Nations Unies ainsi qu'auprès des organisations intergouvernementales et non gouvernementales internationales concernées par la préservation du patrimoine numérique ;
- (b) servir de centre de référence et de forum dans lequel puissent se réunir les Etats membres, les organisations intergouvernementales et non gouvernementales internationales, la société civile et le secteur privé pour élaborer ensemble des objectifs, des politiques et des projets en faveur de la préservation du patrimoine numérique ;
- (c) encourager la coopération, la prise de conscience et la création de compétences, et mettre au point une série de directives éthiques, juridiques et techniques normatives qui serve de manuel de référence complémentaire de la Charte ;
- (d) déterminer, en fonction de l'expérience qui sera acquise au cours des six prochaines années dans l'exécution de la Charte et des Directives présentes, les besoins exacts en instruments normatifs supplémentaires pour encourager et préserver le patrimoine numérique.

## Chapitre 3. Guide des directives

### INTRODUCTION

#### 3.1 L'objectif

Les Directives ont été conçues pour s'adresser à un certain nombre de publics différents et pour couvrir un territoire d'information très vaste. Ce chapitre vise à servir de carnet de route, en aidant les lecteurs à trouver la manière d'utiliser les Directives qui leur conviendra le mieux. (La table des matières, l'index et les notes de renvois figurant en fin de chapitre ont des objectifs similaires.)

#### 3.2 Les publics

Le processus de consultation a révélé au moins quatre publics susceptibles d'utiliser les directives, chacun avec des besoins différents mais qui se rejoignent.

**Les acteurs politique** qui exigent un très haut niveau d'information pour traiter le problème de la préservation numérique, et des données suffisantes pour structurer leur engagement politique en la matière.

Les Directives répondent à ces besoins par :

- l'insertion du projet de charte sur la préservation du patrimoine numérique dans le chapitre 2 ;
- le résumé des principes dans le chapitre 5 ;
- les résumés *En bref* au début de la majorité des chapitres.

**Les responsables de haut niveau** qui cherchent à comprendre les fondements conceptuels de la préservation numérique et les problèmes de gestion auxquels leurs programmes feront face.

Les Directives répondent à ces besoins par :

- les chapitres de la 2e partie qui sont tous axés sur la gestion ;
- les parties consacrées aux *grands défis de la gestion* et aux *principes* des chapitres plus détaillés traitant des méthodes qui se trouvent dans la 3e partie ;
- le résumé des principes dans le chapitre 5.

**Les responsables d'une ligne sur l'Internet** confrontés à des prises de décisions quotidiennes et qui ont besoin à la fois d'une bonne compréhension des notions et d'une idée claire des points détaillés qu'ils auront à gérer.

Les Directives répondent à ces besoins par :

- les chapitres qui présentent une vue d'ensemble des notions dans la 2e partie (notamment les chapitres 7, 8 et 10) ;
- les chapitres détaillés de la 3e partie qui traitent tous de problèmes liés à des méthodes particulières.

*Les utilisateurs techniques* qui ont besoin d'indications techniques précises ainsi que d'une bonne perspective de la manière dont les divers points et processus techniques s'articulent pour fabriquer un programme intégré avec des objectifs de préservation cohérents.

Les Directives n'essaient pas de répondre aux besoins d'informations techniques détaillées qui sont à la fois trop spécifiques d'une situation et trop rapidement dépassés pour figurer dans ces directives. Toutefois, il est recommandé que l'UNESCO crée une section d'Information technique de la version Web de ces Directives où l'on puisse trouver des normes techniques, des manuels et des astuces utiles.

Les Directives doivent, cependant, fournir aux utilisateurs techniques des perspectives d'intégration grâce à l'ordonnance des chapitres. La liste des ouvrages à lire doit aussi fournir un guide utile pour une étude plus approfondie.

### **3.3 Le contenu**

L'ordre des chapitres suit une logique précise.

*La première partie* comprend les éléments d'introduction, notamment l'argumentaire en faveur de la préservation numérique constitué par le projet de Charte de l'UNESCO (chapitre 2); une note sur la terminologie qu'il faut lire avant d'entreprendre la lecture des Directives (chapitre 4) ; et un résumé des principes (chapitre 5).

*La 2e partie* met en perspective les problèmes de gestion. Elle commence par expliquer en quoi consiste le *patrimoine numérique* et explique pourquoi il est menacé (chapitre 6), puis elle présente la *préservation numérique* (chapitre 7), la nature des *programmes de préservation numérique* (chapitre 8), les bases nécessaires pour choisir le type de *responsabilités de préservation* à prendre (chapitre 9), la *gestion des programmes de préservation* (chapitre 10), et les occasions de *travailler en coopération* (chapitre 11).

*La 3e partie* offre un point de vue plus détaillé et centré sur l'aspect opérationnel en étudiant tour à tour chacun des principaux domaines de responsabilités de la gestion quant à la préservation du patrimoine numérique : à commencer par la *sélection de ce qui est suffisamment important pour être conservé* (chapitre 12), puis *travailler avec les producteurs* du patrimoine numérique (chapitre 13), *prendre le contrôle* des matériaux - les transférer, les identifier et les décrire (chapitre 14), *gérer les questions de droits* (chapitre 15), veiller à la *protection de l'authenticité et des données* (chapitre 16), et trouver des solutions pour *garantir les moyens d'accès* (chapitre 17) - point névralgique de la situation d'incertitude qui pèse sur la préservation. Ce chapitre a une structure différente des autres car il cherche à faire une étude comparative d'une série d'options.

Cette partie s'achève sur la proposition de quelques *points de départ de programmes* conçus pour susciter la discussion et la réflexion et sur une série de *prévisions minimales* pour les programmes qui cherchent à lancer une action, quelle qu'elle soit, de préservation numérique (chapitre 18).

*La 4e partie* comporte un glossaire sélectif de termes et une longue liste bibliographique, ainsi que des renvois à des sources fiables pour rester à jour.

### **3.4 Pour les programmes disposant de peu de moyens**

Les Directives ont le souci de proposer des conseils aux personnes qui s'efforcent de créer des programmes avec des moyens extrêmement limités. Tous les chapitres de la 3e partie comportent des propositions prévues à cet effet.

### **3.4 Etudes de cas**

Quelques chapitres de la 3e partie comportent de courtes études de cas. Il s'agit presque toujours de cas fictifs mais fondés sur des expériences réelles. Ce caractère fictionnel a l'avantage de faire ressortir certains aspects pour illustrer un point particulier sans gauchir les programmes réels sur lesquels ils sont éventuellement fondés.

## Chapitre 4. Note sur la terminologie

### INTRODUCTION

#### 4.1 L'objectif

Quelques termes sont employés dans les Directives dans un sens qui peut différer de l'usage habituel. Comme ces termes essentiels reviennent à diverses reprises, il est indispensable d'expliquer d'emblée la signification qui leur est donnée.

D'autres termes, dont l'emploi est moins spécifique, sont expliqués dans le glossaire de la 4<sup>e</sup> partie.

#### 4.2 Les termes

Le terme *préservation numérique* est employé pour décrire les processus impliqués dans la conservation de l'information et d'autres sortes de patrimoine qui existent sous forme numérique. Dans ces directives, il ne renvoie *pas* à l'emploi de l'imagerie numérique ou de techniques de saisie pour faire des copies d'éléments non numériques, même si cela est fait à des fins de préservation. Bien entendu, la copie numérique (également appelée numérisation) peut produire à son tour des matériaux de patrimoine numérique qui exigent d'être préservés.

Le terme *matériaux numériques* est généralement employé comme un terme privilégié pour désigner sur un plan général l'ensemble des éléments du patrimoine numérique. Les termes *objet numérique* ou *ressource numérique* apparaissent aussi à certains endroits. Ces termes ont été employés de manière interchangeable et générique : ils ne renvoient pas à un élément particulier, sauf si c'est clairement spécifié.

Le terme *programme de préservation* sert à désigner tout l'ensemble de dispositifs cohérents qui vise à préserver les matériaux numériques.

On a évité des termes d'usage plus courant comme *archives numériques* ou *gisement de données numérique* en raison des risques d'ambiguïté : le mot *archive* a des acceptions différentes pour la communauté de la gestion de pièces d'archives et la communauté ICT, et les termes respectifs d'*archive* et de *gisement de données* peuvent signifier dans chaque cas qu'il existe un seul site de stockage - ils ne sont donc pas adéquats là où se trouvent des dispositifs très décentralisés.

Bien entendu le terme *programme* n'est pas sans avoir diverses significations. Il doit être compris ici comme englobant tous les aspects de la responsabilité de la préservation, y compris la politique et la stratégie, et sans oublier l'exécution.

Les termes *présentation*, *représentation* sont employés pour décrire les procédés qui permettent l'accès aux matériaux numériques. Le trait d'union du second terme répond à un souci de logique interne, pour bien faire comprendre que la préservation numérique cherche à représenter ce qui a été présenté antérieurement.

## Chapitre 5. Résumé des principes

### INTRODUCTION

#### 5.1 Les objectifs

Ce chapitre vise à regrouper les principales déclarations de principes présentes dans l'ensemble des Directives pour en offrir un résumé aux responsables.

#### 5.2 Les principes

##### 5.2.1 *Le patrimoine*

1. Il n'est pas nécessaire de garder tous les matériaux numériques. Il faut conserver seulement ceux que l'on estime avoir une valeur durable : ce sont ceux-ci qui constituent le patrimoine numérique.

2. Pour ces matériaux jugés dignes d'être gardés, le maintien continu de leur survie et de leur accessibilité est dans une phase critique. Quand l'accès à de grandes quantités de données a été perdu, les chances de le retrouver deviennent très minces. Le maintien continu de cet accès réclame une action soutenue, directe (on l'appelle *préservation numérique*) ; l'"attention bienveillante" est une attitude passive qui est loin de suffire.

##### 5.2.2 *La préservation numérique*

3. Il est impossible d'affirmer que les matériaux numériques sont préservés si leur accès est perdu. La préservation a pour objectif de maintenir la capacité de présenter l'essentiel des matériaux numériques authentiques.

4. La préservation numérique doit prévenir les dangers qui menacent sur toutes les couches de l'objet numérique : physiques, logiques, conceptuelles et essentielles.

##### 5.2.3 *La responsabilité*

5. La préservation numérique n'est possible que si des organisations et des individus en endossent la responsabilité. Le point de départ de toute action est de prendre une décision quant à cette responsabilité.

6. Chacun ne doit pas tout faire et tout ne doit pas être fait en même temps.

7. Des programmes complets et fiables sont fort souhaitables, mais ils ne peuvent être réalisés dans toutes les situations où il y a un besoin. Le cas échéant, mieux vaut, en règle générale, prendre des mesures non limitées et incertaines plutôt que pas de mesure du tout. Il est préférable habituellement d'avancer à petits pas plutôt que de rester sur place.

8. En prenant des mesures, les responsables doivent être conscients que des problèmes complexes sont en jeu. Il est important de ne causer aucun dommage. Aussi doivent-ils s'efforcer de comprendre l'ensemble du processus et la finalité des objectifs à atteindre pour éviter les initiatives qui compromettront une action de préservation ultérieure.

9. La prise de responsabilité doit être affirmée de façon explicite et en toute connaissance de cause, en tenant compte des implications probables qui en résulteront pour d'autres programmes de préservation et les autres parties prenantes.

#### **5.2.4 Décider ce qu'il faut garder**

10. Les décisions en matière de sélection doivent être bien étayées, cohérentes et justifiables.

11. La décision de préserver peut faire l'objet d'une révision ultérieure ; la décision de ne pas préserver est habituellement définitive.

#### **5.2.5 Travailler avec les producteurs**

12. Actuellement, les efforts de préservation doivent lutter contre la tendance prédominante de la technologie numérique et la manière dont elle est développée et utilisée.

13. Les matériaux numériques sont souvent créés sans envisager une préservation à long terme.

14. Travailler en concertation avec les producteurs pour influencer sur les standards et les techniques qu'ils utilisent, et pour les rendre plus sensibles aux nécessités de la préservation sont des investissements importants.

#### **5.2.6 Les droits**

15. Les programmes de préservation doivent clarifier leurs droits juridiques pour rassembler, copier, nommer, modifier, préserver et fournir l'accès aux matériaux numériques qu'ils prennent en charge.

#### **5.2.7 Le contrôle**

16. Les matériaux du patrimoine numérique doivent être placés en un lieu sûr où on peut les contrôler, les protéger et les administrer en vue de leur préservation.

17. Les matériaux du patrimoine numérique doivent être exclusivement identifiés et décrits en utilisant des métadonnées adaptées à la découverte des ressources, à la gestion et à la préservation.

18. Agir ensuite à bon escient dépend d'une documentation adéquate. Il est plus facile de documenter les caractéristiques des ressources numériques proches de leur source que d'établir cette documentation après coup.

19. Les programmes de préservation doivent utiliser des schémas de métadonnées normalisés au fur et à mesure qu'ils deviennent disponibles pour la compatibilité entre programmes.

20. Il faut soigneusement entretenir les liens entre les objets numériques et leurs métadonnées, de même qu'il faut préserver les métadonnées.

#### **5.2.8 L'authenticité et la protection des données**

21. L'authenticité est un problème crucial lorsque les objets numériques servent de preuves. Elle peut également être importante pour d'autres sortes de patrimoine numérique.

22. Les données qui sont à la base des objets numériques doivent être soigneusement sauvegardées et gérées s'il y a une quelconque chance de représenter les objets authentiques aux utilisateurs.

23. Les programmes de préservation numérique suscitent une préoccupation d'authenticité grandissante parce qu'ils doivent très souvent recourir à des procédés qui impliquent un changement.

24. Des mesures qui garantissent que l'intégrité des données n'est pas mise en péril et une documentation qui conserve l'identité lisible du matériau, tels sont les plus sûrs moyens de protéger l'authenticité.

25. La protection des données repose sur les principes de sécurité du système et sur la redondance. Pour les programmes de préservation, la redondance peut comporter des sauvegardes soigneusement stockées qui sont conçues autour d'un entretien à long terme des données plutôt qu'un cycle d'écrasement d'anciennes données par de nouvelles.

### **5.2.9 Maintenir l'accessibilité**

26. Pour parvenir à une accessibilité permanente, il faut trouver des moyens compétitifs de garantir l'accès chaque fois qu'il est nécessaire, autant à court terme qu'à long terme.

27. Les normes sont des bases importantes pour la préservation numérique, mais de nombreux programmes doivent trouver les moyens de préserver l'accès à des matériaux mal normalisés, dans un environnement de normes changeantes.

28. L'acte de préserver ne devrait pas être retardé jusqu'à l'apparition de "normes de préservation numériques" uniques.

29. Les données numériques sont toujours tributaires d'une combinaison d'outils logiciels et matériels pour la consultation, mais le degré de dépendance par rapport à des outils spécifiques détermine la gamme des options de préservation.

30. Il est raisonnable que les programmes choisissent des stratégies multiples pour préserver l'accès, notamment à des collections diverses. Ils doivent prendre en compte les avantages potentiels de la conservation des trains de données originaux ainsi que de toutes les versions modifiées pour s'assurer contre l'échec de stratégies encore peu sûres.

31. Les stratégies, pour préserver l'accessibilité, ne sont pas isolées : elles s'appuient sur d'autres responsabilités, et l'association de nombreuses stratégies peut se révéler efficace.

32. Les programmes de préservation sont souvent nécessaires pour juger acceptables ou inacceptables les niveaux de perte, en ce qui concerne les articles, les éléments et les besoins de l'utilisateur.

### **5.2.10 La gestion**

33. Attendre qu'apparaissent des solutions exhaustives et fiables avant de prendre une mesure responsable signifiera probablement la perte du matériau.

34. Les programmes de préservation réclament une gestion avisée. Celle-ci réside en grande partie dans des compétences générales en la matière conjuguées à une connaissance suffisante des problèmes de préservation numérique pour prendre les bonnes décisions au moment voulu.

35. La préservation numérique comprend l'évaluation et la gestion des risques.

36. Les programmes se heurtent généralement à une trop grande masse de matériaux et de problèmes pour pouvoir y faire face, aussi doivent-ils fixer des priorités.

37. Les coûts des programmes de préservation sont difficiles à évaluer parce qu'ils présentent beaucoup d'incertitudes, notamment l'évolution des techniques, le renouvellement des technologies et des calendriers très étalés. Les coûts peuvent être plus bas par unité d'information que pour des matériaux non numériques, mais le volume d'information à gérer sous forme numérique est considérable, de sorte que les coûts totaux resteront probablement élevés, y compris les coûts de configuration et les frais courants importants.

38. Les programmes de préservation peuvent démarrer en tant que projets pilotes, mais en fin de compte, ils ont besoin de créer des modèles de gestion durables.

39. L'on peut trouver des prestataires de service appropriés pour remplir certaines de ces fonctions, mais l'ultime responsabilité de la mise en oeuvre de la préservation repose sur les programmes de préservation et sur ceux qui les surveillent et les ressource.

#### ***5.2.11 Travailler ensemble***

40. Travailler en collaboration est souvent une manière rentable de créer des programmes de préservation avec une couverture large, un soutien mutuel et les compétences requises.

41. La collaboration implique des coûts et des choix autant que des avantages potentiels.

**DEUXIEME PARTIE**  
**PERSPECTIVES DE GESTION**

## Chapitre 6. Comprendre le patrimoine digital

### INTRODUCTION

#### 6.1 Les objectifs

Le propos de ce chapitre est de présenter les concepts de *patrimoine numérique* et de *continuité numérique*. Il vise à aider les lecteurs à comprendre la valeur et la portée des matériaux du patrimoine numérique et les menaces qui pèsent sur leur survie. Ce sont des notions essentielles pour les responsables comme pour ceux qui conçoivent et réalisent les programmes.

#### 6.2 En bref

Le *patrimoine numérique* est constitué de matériaux fondés sur l'informatique, d'une valeur durable, qu'il est nécessaire de conserver pour les générations futures. Il émane de communautés, de régions, d'industries et de secteurs différents. Tous les matériaux numériques ne sont pas de valeur durable, mais ceux qui le sont exigent des méthodes actives de préservation si l'on veut que la continuité du patrimoine numérique soit assurée.

### PERSPECTIVE DE GESTION

#### 6.3 Patrimoine et patrimoine numérique

Le patrimoine est défini par les documents de l'UNESCO comme étant "notre héritage du passé, ce avec quoi nous vivons aujourd'hui, et ce que nous transmettons aux générations futures". Un patrimoine est quelque chose qui est, ou devrait être, transmis de génération en génération parce qu'il est précieux.

L'idée de patrimoine culturel est familière : ce sont des sites, des objets et des choses immatérielles qui ont une valeur culturelle, historique, esthétique, archéologique, scientifique, ethnologique ou anthropologique pour des groupes et des individus. La notion de patrimoine naturel est également très familière : elle recouvre des caractéristiques physiques, biologiques et géologiques ; des habitats de plantes ou d'espèces animales et des domaines précieux sur le plan scientifique ou esthétique ou du point de vue de la conservation.

Y a-t-il un *patrimoine numérique* émergent ?

Selon le projet de charte pour la préservation du patrimoine numérique :

- les ressources des connaissances et de l'expression humaines, qu'elles soient culturelles, éducatives, scientifiques et administratives, ou qu'elles englobent des informations techniques, juridiques, médicales et autres, sont de plus en plus créées de façon numérique, ou converties sous forme numérique à partir de ressources analogiques existantes. Là où les ressources sont d'"origine numérique", il n'existe pas d'autre format que l'original numérique ;
- les matériaux numériques comprennent des textes, des bases de données, des images fixes et animées, du son, des graphiques, des logiciels et des pages Web sous une grande et

croissante variété de formats. Ils sont souvent éphémères et exigent une production, un entretien et une gestion avisés ;

- nombre de ces ressources sont d'une valeur et d'une importance durables et constituent donc un patrimoine à protéger et à préserver pour les générations actuelles et futures. Ce patrimoine peut exister dans n'importe quelle langue, n'importe quelle partie du monde et dans tous les domaines de connaissances ou d'expression humaines.

En utilisant des ordinateurs et des outils connexes, les hommes créent et partagent des ressources numériques - informations, expression créatrice, idées et connaissances codées pour le traitement informatique - auxquelles ils attachent de la valeur et qu'ils veulent partager avec d'autres dans le temps autant que dans l'espace. Cela prouve qu'il existe un patrimoine numérique. C'est un patrimoine constitué de nombreuses parties, qui partage de nombreuses caractéristiques communes et qui est exposé à de nombreuses menaces communes.

Les définitions de patrimoine doivent être replacées chaque fois dans leur contexte. Par exemple, l'UNESCO définit un *patrimoine mondial* comme étant constitué de sites de valeur culturelle et naturelle remarquables au plan universel, qu'il faut préserver ; de nombreuses juridictions nationales et étatiques définissent également leur propre patrimoine national, régional ou d'Etat. Cependant, la valeur patrimoniale peut aussi reposer sur ce qui est important à l'échelle d'un groupe ou d'une communauté. Les biens patrimoniaux peuvent exister hors des divers cadres proposés par une juridiction nationale ou des conventions internationales. Tout ce qui est considéré comme assez important pour être transmis à la postérité est susceptible de détenir une valeur patrimoniale.

Il est probable que ce patrimoine numérique s'agrandira et se répandra avec le temps. De plus en plus, les particuliers, les organisations et les communautés utilisent les technologies numériques pour documenter et exprimer ce à quoi ils attachent de la valeur, et ce qu'ils veulent transmettre aux générations futures. Ont émergé de nouvelles formes de communication et d'expression qui n'existaient pas auparavant. L'Internet est un vaste exemple de ce phénomène.

Il est également probable que le développement d'outils pour permettre une utilisation en plusieurs langues et écritures de l'Internet conduira à son augmentation rapide dans certaines parties du monde, actuellement désavantagées par l'emploi prédominant de l'anglais sur l'Internet.

Veiller à ce que ce patrimoine numérique foisonnant demeure disponible est donc un thème universel qui concerne tous les pays et toutes les communautés.

#### **6.4 Les différents types de matériaux numériques**

Avec les années, on peut s'attendre qu'émergent de nouveaux types de patrimoine numérique : la puissance de base de la technologie a déjà donné naissance à des formes aussi diverses que le traitement de texte, le courriel, les sites Web, les bases de données relationnelles, les modèles et les simulations informatiques, le son et la vidéo numériques, l'imagerie spatiale et les jeux électroniques. A l'heure où l'on écrit ces lignes, le patrimoine numérique présente une large palette de matériaux, entre autres (cette liste n'est nullement exhaustive) :

- les publications électroniques, qui s'adressent à un large public. Elles sont diffusées de diverses manières, y compris en direct, via la "toile mondiale", ou sur des supports portatifs comme les CD-ROM, les DVD, les disquettes souples et divers dispositifs du livre électronique. Certaines parviennent à conjuguer l'accès du réseau et du support portatif à différentes parties de la publication. De même que leurs moyens de diffusion, on peut classer les publications numériques par genre : certains nous sont familiers car ils viennent des formats de l'édition traditionnelle, telles les monographies et les séries, alors que d'autres sont moins faciles à définir comme les sites Web et les revues électroniques. Certaines publications sortent sous forme de produit fini, mais d'autres évoluent avec le temps, leurs créateurs tirant parti du potentiel interactif de l'Internet. L'édition de l'imprimé continue à augmenter, mais un nombre croissant de publications paraît dans des versions numériques et de plus en plus dans des versions exclusivement numériques. Les éditeurs commerciaux et les éditeurs non commerciaux produisent des publications numériques, ainsi que le font des millions d'autres personnes qui ne se considèrent aucunement comme des éditeurs ;
- des matériaux semi-édités comme les communications universitaires préimprimées et les thèses conservées en pré tirage électronique, ainsi que d'autres archives destinées à un usage restreint au sein de communautés spécifiques comme les universités et les sociétés savantes ;
- des rapports d'activité organisationnels et personnels, des transactions, des correspondances, etc. Une très grande partie des archives commerciales et gouvernementales mondiales existe désormais dans des systèmes de conservation d'archives électroniques. Le courriel, les messages adressés aux listes de discussion et aux microserveurs, les agendas, les journaux intimes et les caméras du Web - des interactions informelles et dynamiques rendues possibles par la technologie numérique - peuvent aussi comporter d'importantes pièces d'archives numériques au milieu d'un raz de marée de données ;
- des ensembles de données pour enregistrer et analyser les phénomènes scientifiques, géospatiaux, spatiaux, sociologiques, démographiques, éducatifs, sanitaires, environnementaux et autres ;
- des objets pédagogiques employés dans l'éducation assistée par la technologie ;
- des outils logiciels comme les bases de données, les modèles, les simulations et les applications logicielles ;
- des matériaux uniques inédits qui peuvent comporter des rapports de recherche, de l'histoire orale et des témoignages de culture populaire ;
- des "manuscrits" électroniques comme les brouillons d'oeuvres et les correspondances personnelles ;
- des produits récréatifs issus des industries du film, de la musique, de la radiodiffusion et des jeux, à la fois commerciaux et non commerciaux ;
- des illustrations et des photographies documentaires créées par numérisation ;
- des copies numériques d'images, de son, de texte et d'objets en trois dimensions à partir d'originaux non numériques.

Un grand nombre de ces matériaux n'existent que sous forme numérique (même s'ils sont transportés sur un quelconque support physique). En l'absence d'une version non numérique équivalente, leur contenu est particulièrement exposé aux atteintes qui menacent les matériaux numériques.

Il existe aussi des collections de copies numériques dont le nombre connaît une croissance rapide. Créées à l'origine à partir de sources non numériques, elles peuvent paraître moins vulnérables, mais beaucoup d'entre elles restent en fait la seule source existante d'originaux qui ont depuis été détériorés, perdus ou dispersés.

## **6.5 La continuité numérique**

La continuité du patrimoine numérique est d'une extrême importance. De plus en plus, c'est un patrimoine qui apporte des documents sur l'action des gouvernements, les résultats de la recherche scientifique, le débat d'idées, les aspirations et les rêves des communautés, l'histoire du monde actuel et à venir.

Pour que ces documents ne soient ni perdus ni dénaturés, la continuité est nécessaire : la continuité de production, de survivance et d'accès. Il faut réaliser cette triple continuité pour parer à plusieurs menaces :

- les supports utilisés pour stocker ces matériaux numériques sont habituellement instables et se détériorent en l'espace de quelques années ou décennies tout au plus ;
- l'utilisation des matériaux numériques dépend de moyens d'accès qui fonctionnent de manière particulière : par une combinaison d'outils souvent complexes, comprenant un matériel et des logiciels qui deviennent, de manière caractéristique, périmés en quelques années et sont remplacés par de nouveaux outils qui fonctionnent différemment ;
- les matériaux peuvent être perdus lors de catastrophes comme les incendies, les inondations, les défaillances de matériel ou une attaque directe ou de virus qui rendent inutilisables les données conservées et les systèmes de fonctionnement ;
- les barrières d'accès, comme la protection par mot de passe, le cryptage, les dispositifs de sécurité ou les chemins d'accès à code renforcé, peuvent empêcher l'accès continu au-delà des circonstances très limitées pour lesquelles elles ont été conçues ;
- la valeur du matériau risque de n'être pas reconnue avant qu'il soit perdu ou transformé ;
- personne ne peut prendre la responsabilité du matériau même si sa valeur est reconnue ;
- ceux qui prennent la responsabilité des matériaux n'ont pas les connaissances et les compétences voulues ;
- les moyens disponibles se révèlent insuffisants pour soutenir l'action de préservation sur la période donnée ;
- les autorisations juridiques nécessaires à la préservation peuvent ne pas être obtenues ;
- le temps ou les compétences disponibles manquent pour répondre suffisamment vite à un changement important et brutal de technologie ;

- il arrive que les matériaux numériques soient bien protégés, mais si mal identifiés et décrits, que les utilisateurs potentiels sont dans l'impossibilité de les trouver ;
- il peut y avoir une telle perte d'informations contextuelles que les matériaux eux-mêmes sont inintelligibles ou non fiables même quand on peut y accéder ;
- des aspects fonctionnels d'un rôle décisif, comme le formatage des documents ou les règles selon lesquelles fonctionnent les bases de données, faute d'avoir pu être identifiés, sont mis au rebut ou endommagés au cours du processus de préservation.

Les mesures à prendre pour assurer la continuité face à ces menaces particulières et à d'autres encore sont englobées sous le terme de *préservation numérique* - une nouvelle forme de préservation qui se rapporte spécifiquement aux matériaux du patrimoine numérique.

## CONSIDERATIONS PARTICULIERES

### 6.6 La stabilité de l'Internet comme scénario de risque spécifique

L'Internet offre un cas de figure intéressant dans lequel nombre de ces menaces entrent en ligne de compte. En évaluant les risques associés à l'Internet, il faut veiller à bien faire la distinction entre deux points de vue qui se rejoignent mais n'en sont pas moins différents.

L'un consiste à considérer l'Internet comme un tout. Or il n'y a pas d'agence centrale à pouvoir décider ce qu'il arrive au matériau qui devient disponible sur l'Internet. Les utilisateurs se connectent à un kaléidoscope d'informations dont aucun créateur, éditeur ou autre agence n'est responsable - c'est là un trait spécifique de la Toile. Faut-il laisser se perdre ce pot-pourri d'informations et l'expérience de sa pratique sous prétexte que personne n'en est responsable ?

*L'examen des ressources individuelles publiées sur l'Internet apporte un autre éclairage. Elles peuvent être également tout à fait instables, mais cette instabilité reflète les décisions et les actions de leurs propriétaires lorsqu'ils les effacent, les modifient, les déplacent ou leur donnent un nouveau nom. La perte de matériaux qui se produit dans cet environnement est attribuable à la gestion dont ils ont fait l'objet au niveau local étant donné qu'ils sont exposés à de nombreuses menaces semblables à celles qui pèsent sur d'autres sortes de matériaux numériques.*

Contrairement aux utilisateurs de l'Internet, ceux qui possèdent et gèrent les objets et les sites numériques ont le pouvoir de vérifier si l'information reste valable. Si on les charge de maintenir l'accès, c'est qu'ils sont généralement en mesure de le faire.

Toutefois, l'Internet présente des risques particuliers. Par exemple :

- il y a un facteur de nouveauté important : des éditeurs choisissent souvent de changer l'ordre des choses - tantôt la manière dont est présentée l'information, tantôt le contenu ;
- un grand nombre de ressources de l'Internet sont faites d'un matériau composite d'éléments virtuels de sources diverses qui ne peut pas être conservé dans son ensemble n'importe où. En modifiant un élément, on risque de détruire l'ensemble ;
- le sentiment d'accès planétaire peut conduire certains responsables de l'information à se croire capables de recréer leur information si elle est perdue, sans savoir que celle-ci existe dans un système local et qu'elle est vulnérable aux détériorations et aux pertes associées à ce système. Les responsables de l'information risquent de négliger de prendre les mesures

de sauvegarde et de sécurité normales comme ils l'auraient fait automatiquement dans un système autonome ;

- il est possible de publier des matériaux numériques sur l'Internet avec une grande facilité et à très bas prix. Du coup, de nombreux "éditeurs" ne prévoient pas de conserver leurs publications ou n'en ont pas les moyens : leurs oeuvres sont alors véritablement éphémères.

## LES ETUDES DE CAS

(Ces études de cas fictionnels ont été choisies pour illustrer quelques exemples de matériaux du patrimoine numérique - et non pas nécessairement pour avaliser leur mode de gestion.)

Un service gouvernemental a récemment mis en circulation des ordinateurs à l'usage de chacun des membres de son personnel de sorte qu'il puissent rédiger leurs propres lettres, mémos internes, rapports et envoyer des messages par courriel. Une directive émise par le service stipule que tous les documents finalisés et les projets importants ainsi que les courriels professionnels doivent être archivés en vue d'être conservés à long terme. (Ils font partie du patrimoine numérique.) Cependant, les courriels personnels et les brouillons n'ont pas besoin d'être conservés.

Une communauté rurale isolée se préoccupe depuis longtemps du fait que sa cohésion traditionnelle est en train de se perdre en même temps que le respect de son mode de vie. Les anciens décident d'enregistrer tout ce qu'ils peuvent sur les traditions de la communauté et utilisent un réseau informatique pour l'enregistrer et le partager. Celui-ci devient un véritable pôle d'aimantation qui renouvelle l'intérêt et la fierté de vivre dans la communauté pour la grande majorité de ses membres, et se révèle également une source de revenus collective puisqu'un choix de bases de données est mis à disposition d'étrangers autorisés. La communauté est tombée d'accord pour conserver et gérer elle-même les bases de données en nombre croissant.

Un studio d'enregistrement réalise des enregistrements numériques depuis dix ans. Les originaux sont stockés sur un grand nombre de bandes et de CD-ROM qui servent parfois à regrouper les nouveaux produits destinés aux sociétés d'enregistrement locales, mais qui sont le plus souvent simplement laissés en archivage. Tous les 12 mois, le responsable les trie et jette toutes les vieilles bandes qui ne lui paraissent pas intéressantes ou qui semblent en bout de course.

Une professeur d'université installe un site Web pour encourager la discussion au sein de sa discipline. Elle participe à des conférences et elle publie régulièrement des articles dans des revues spécialisées, mais découvre que le débat de loin le plus intéressant dans son domaine a lieu sur son agenda Web rattaché à son site Web. Elle s'inquiète du fait qu'il va être perdu sans qu'il en reste la moindre trace et que les futurs étudiants et chercheurs n'auront aucune idée de la manière dont certains concepts ont été discutés à l'origine.

Un institut de recherches étudie les variations du débit et des crues dans le système d'un fleuve important, en enregistrant des données exhaustives sur plusieurs décennies et en utilisant des simulations informatiques diverses pour reproduire les effets des différentes précipitations dans les zones de captage. Ainsi se rendent-ils compte que les changements de mode d'exploitation des sols entraînent des modifications dans leurs données.

## REFERENCES - pour en savoir plus

### Renvois

*Les menaces*, voir La gestion du risque, chapitre 10.

## Chapitre 7. Comprendre la préservation numérique

### INTRODUCTION

#### 7.1 Les objectifs

Ce chapitre entend aider ceux à qui peut incomber la responsabilité d'assurer la préservation de matériaux du patrimoine numérique à comprendre la nature fondamentale, les objectifs et les stratégies de ce type de préservation. Ce sont des notions importantes à comprendre tant pour ces responsables que pour ceux qui conçoivent et réalisent les programmes.

#### 7.2 En bref

La *préservation numérique* consiste en un certain nombre de méthodes qui visent à assurer l'accessibilité permanente aux matériaux numériques. Faire cela implique qu'on ait trouvé les moyens de représenter ce qui a été présenté à l'origine aux utilisateurs moyennant une combinaison d'outils logiciels et matériels agissant sur des données. Réussir cela exige que les objets numériques soient compris et gérés sur quatre plans : en tant que phénomènes physiques ; en tant que codages logiques ; en tant qu'objets conceptuels qui ont un sens pour les humains ; et en tant qu'ensembles d'éléments essentiels à préserver en vue d'offrir aux futurs utilisateurs l'essence de l'objet.

### PERSPECTIVE DE GESTION

#### 7.3 La préservation numérique

On peut considérer la préservation numérique comme la somme de tous les procédés qui visent à assurer l'accès permanent des matériaux du patrimoine numérique aussi longtemps que le besoin s'en présente.

Les menaces les plus importantes qui pèsent sur la continuité numérique sont liées à la perte des moyens d'accès. On ne peut affirmer que les matériaux numériques sont préservés si les moyens d'accès ont été perdus et si l'accès devient impossible. La préservation des matériaux numériques a pour but de maintenir l'accessibilité : la capacité d'accéder à leur message, ou à leur propos, essentiel et authentique.

#### 7.4 Aborder la préservation numérique comme une "performance"

Il y a une ressemblance fondamentale entre la manière dont on accède aujourd'hui aux objets numériques et celle dont on y accédera dans toute utilisation future. Dans les deux cas, on peut considérer cet accès comme une "performance"<sup>1</sup>.

---

<sup>1</sup> Cette notion est débattue dans Heslop H. Davis S (2002) (inédit), *An Approach to the Preservation of Digital Records*, Archives nationales d'Australie, Canberra.

Les objets numériques sont rendus accessibles par l'application d'outils logiciels et matériels à des données afin de créer une présentation ou une "performance" qui a un sens pour l'utilisateur. Ce peut être la présentation d'un document de traitement de texte, ou d'un morceau de son enregistré, ou d'une page Web, ou les résultats d'une demande de bases de données, ou de toute autre sorte d'objet numérique qui dépend de la manière dont les données sont codées et des actions pour lesquelles sont programmés les outils. Si nous appliquons les mêmes outils aux mêmes données, nous nous attendons à obtenir à chaque fois la même "performance".

La préservation numérique doit fonctionner de la même manière, en représentant d'une façon ou d'une autre ce que l'on juge être les éléments essentiels de la "performance" originale lorsqu'on en aura besoin ultérieurement.

Ainsi conçue, la préservation numérique apparaît comme simple et directe. En effet, la copie des données d'un support à un autre et la fourniture d'outils adéquats pour recréer la "performance" attendue préserveront la continuité de l'accès à la plupart des objets numériques.

Cependant, ce modèle simple comporte des points forts complexes : il peut être ardu de définir la "performance" qui doit être représentée ; il est généralement difficile d'établir quels outils sont nécessaires une fois que ceux d'origine ont disparu ; les outils eux-mêmes sont tributaires, comme il se doit, d'autres outils qui risquent d'être dépassés ; il peut se révéler également difficile de trouver des outils qui créeront la "performance" demandée d'une manière fiable, à prix compétitif et en temps opportun, en particulier dans un contexte de plusieurs milliers, millions, ou plus, d'objets numériques.

Malgré cette multiple complexité sous-jacente, le modèle de "performance" aide à reconnaître les objectifs que les programmes de préservation numérique doivent se fixer : les meilleurs moyens de représenter ce dont les utilisateurs ont besoin pour disposer d'un accès.

## 7.5 Comprendre les matériaux qui sont préservés

Les programmes de préservation doivent traiter les objets numériques sous quatre formes :

- en tant qu'**objets physiques** qui consistent en "inscriptions" (généralement des états binaires de "marche" ou "arrêt") sur des supports comme les disques ou les bandes informatiques. Bien qu'on ait l'impression qu'elles existent dans le "cyberespace", les ressources en ligne doivent exister quelque part sur des supports tangibles ;
- en tant qu'**objets logiques**, c'est-à-dire en codes informatiques lisibles, dont l'existence à tout moment dépend des inscriptions physiques mais n'est pas liée à un support particulier ;
- en tant qu'**objets notionnels** qui ont un sens pour les humains, contrairement aux objets logiques ou physiques qui les encodent à tout moment (c'est identifiable comme étant la "performance" présentée à l'utilisateur) ;
- comme des lots d'**éléments essentiels** qui concrétisent le message, le propos ou les caractéristiques pour lesquels le matériau a été choisi en vue d'être préservé.

La nature multicouches des objets numériques a des conséquences profondes pour la préservation numérique : celle-ci implique des choses différentes pour chaque couche<sup>2</sup>.

---

<sup>2</sup> Cette notion est adaptée de Thibodeau K. (2002).

Les programmes de préservation du patrimoine non numérique se préoccupent traditionnellement de préserver l'objet physique en tant que concrétisation du sens de l'objet. Or les manifestations physiques individuelles d'un objet numérique sont presque inévitablement perdues les unes après les autres, car les supports utilisés pour le stockage physique sont par nature instables et exposés à une détérioration à court terme. La préservation exige une succession de transferts de données d'un support physique à un autre.

Malgré ce déplacement du pôle d'intérêt de l'objet physique à un objet notionnel inhérent à la préservation numérique, on ne doit jamais oublier que les objets numériques ne peuvent survivre sans quelque forme tangible appropriée.

Le codage logique a normalement une existence bien plus longue que n'importe quelle inscription particulière, mais il n'est nullement intouchable. Puisque les couches technologiques utilisées pour l'accès - le matériel, tels les processeurs informatiques, lecteurs de disquettes et périphériques, ainsi que de nombreuses couches de logiciels tels les systèmes d'exploitation, applications spécifiques et outils de présentation - deviennent dépassées, il peut apparaître nécessaire de changer le codage logique afin qu'il présente le même objet notionnel moyennant une technologie différente.

L'objet notionnel est la préoccupation ultime de la préservation ; comme on l'a signalé plus haut, c'est à ce niveau que les objets numériques transmettent le sens aux utilisateurs humains.

Il existe cependant, pour la plupart des objets numériques, une couche supplémentaire à prendre en compte. Nombre d'entre eux se composent de plusieurs éléments, certains étant plus importants que d'autres dans le transport du message essentiel de l'objet. Les programmes de préservation doivent choisir les sous-ensembles d'éléments à préserver pour la représentation aux utilisateurs.

## **7.6 Des stratégies pour préserver les matériaux numériques**

La préservation numérique exige que l'on choisisse de mettre en oeuvre un registre évolutif de stratégies pour parvenir au type d'accessibilité dont on a parlé plus haut, en traitant les besoins de préservation des différentes couches des objets numériques. Ces stratégies supposent certaines conditions :

- travailler avec les producteurs (créateurs et diffuseurs) afin d'appliquer des normes qui prolongeront la vie effective des moyens d'accès disponibles et réduire la gamme des problèmes inconnus à gérer ;
- admettre qu'il n'est pas productif d'essayer de tout préserver et sélectionner le matériau à préserver ;
- mettre le matériau en lieu sûr ;
- contrôler le matériau, en utilisant des métadonnées structurées et une autre documentation, pour faciliter l'accès et soutenir tout processus de préservation ;
- protéger l'intégrité et l'identité des données ;
- choisir des moyens appropriés de fourniture d'accès en cas de changement technologique ;

- gérer les programmes de préservation pour qu'ils atteignent leurs objectifs dans des conditions qui soient compétitives, opportunes, globales, dynamiques et responsables.

## REFERENCES - pour en savoir plus

### Renvois internes

Certains thèmes et notions présents dans ce chapitre sont examinés plus loin dans ces Directives :

*Eléments essentiels* voir chapitres 12 et 17,

*Stratégies de préservation*, voir les chapitres appropriés dans la 3e partie.

### Renvois externes (tous établis en mars 2003)

- Heslop H., Davis S. (2002) (inédit). *An Approach to the Preservation of Digital Records*. Archives nationales d'Australie, Canberra.
- Thibodeau K. (2002). Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years. In: *The State of Digital Preservation: An International Perspective - Conference Proceedings*, Documentation Abstracts, Inc., Institutes for Information Science, Washington, D.C., April 24-25, 2002. Conseil sur les ressources des bibliothèques et de l'information, Washington, D.C. <http://www.clir.org/pubs/reports/pub107/thibodeau.html>

## Chapitre 8. Comprendre les programmes de préservation numérique

### INTRODUCTION

#### 8.1 Les objectifs

Dans ce chapitre, le lecteur trouvera des informations de haut niveau sur les responsabilités, les fonctions et les caractéristiques de programmes de préservation numériques complets et fiables. Ces informations sont importantes pour les responsables comme pour ceux qui conçoivent les programmes.

#### 8.2 En bref

Les programmes de préservation comprennent un certain nombre de responsabilités et de fonctions qu'on a déjà définies, du moins à un niveau conceptuel. Des programmes exhaustifs doivent prendre en charge des matériaux numériques appropriés et garantir qu'ils restent compréhensibles et utilisables en tant que copies authentiques. Ceci implique généralement de mettre les matériaux, dûment préparés, en même temps qu'une documentation ou des métadonnées correspondantes, dans un système d'archivage numérique, où l'on peut gérer les menaces de perte de données et de changement technologique. Les caractéristiques ou les attributs des programmes sur lesquels on peut compter pour assurer une préservation numérique continue ont également été décrits sous divers angles : responsabilité, viabilité, durabilité, compatibilité technique, sécurité et justification.

### PERSPECTIVE DE GESTION

#### 8.3 Quelques notions

##### 8.3.1 Les programmes de préservation

Dans ces directives, les ensembles de dispositifs mis en place pour rendre effective la préservation numérique sont appelés *programmes de préservation*. C'est une notion large qui inclut autant la politique que les aspects pratiques de la mise en oeuvre.

##### 8.3.2 Les lieux sûrs

Ces Directives supposent que les matériaux du patrimoine numérique doivent être déplacés d'un environnement opératoire dans un *lieu sûr*, autrement dit des archives. Là, on peut les abriter des dangers qu'ils courent sur les plans physique et logique, et les gérer en vue d'une accessibilité continue.

(Selon un argument contraire, les matériaux numériques auraient une bien plus grande chance de survivre s'ils restaient fréquemment utilisés, car l'on s'efforcerait alors de les garder accessibles. Le matériau placé dans des archives "mortes" risque davantage d'être négligé et de ne pas bénéficier d'une action de préservation avant qu'il ne soit trop tard. Cet argument met l'accent sur deux vérités importantes : il ne faut pas négliger l'action préservatrice ; le matériau que l'on réclame a plus de chances de survivre que le matériau qui n'est pas utilisé. Mais cet argument ne s'applique pas aux matériaux du patrimoine, qu'il convient de préserver même s'ils ne font souvent l'objet que d'une faible demande d'utilisation. Que le matériau soit souvent utilisé ou non, il doit exister une copie

qu'il faudra soigneusement stocker et gérer au cas où il doit survivre, même si cela implique la création d'un lieu sûr au sein d'un environnement de travail.)

### **8.3.3 Les ensembles cohérents d'informations**

Les objets d'information numérique ne sont généralement pas compréhensibles ou représentables en eux-mêmes : les utilisateurs ont besoin d'aide pour s'en servir. La préservation dépend de la conservation des objets numériques et de tous les outils et informations permettant d'y accéder et de les comprendre. Réunis, ils composent ce qu'on peut appeler un *ensemble cohérent d'informations* qui doit être géré soit comme un objet unique soit comme un ensemble virtuel (avec l'objet et les outils d'informations associés qui sont reliés mais stockés séparément).

## **8.4 Les responsabilités des programmes de préservation exhaustifs**

Les programmes de préservation qui visent à être exhaustifs ont les responsabilités suivantes :

- négocier et se procurer les matériaux numériques appropriés auprès des producteurs ;
- contrôler suffisamment le matériau pour assurer sa préservation sur le long terme ;
- définir à qui s'adresse le matériau conservé et qui aura besoin de pouvoir le comprendre ;
- garantir que le matériau restera compréhensible pour le groupe d'utilisateurs prévus ;
- garantir que le matériau est protégé contre toute menace éventuelle et faire en sorte qu'il soit disponible et que son authenticité soit garantie ;
- rendre le matériau disponible pour le groupe d'utilisateurs ciblé ;
- préconiser une bonne pratique de la création de ressources numériques.

## **8.5 Les fonctions des programmes de préservation exhaustifs**

Pour remplir ces responsabilités, les programmes de préservation qui cherchent à être exhaustifs doivent s'acquitter des fonctions suivantes :

### **8.5.1 Créer ou trouver un lieu sûr**

Les programmes de préservation doivent trouver un lieu sûr où les matériaux numériques puissent être stockés et gérés. Etant donné que le concept du programme de préservation permet la répartition des dispositifs et le partage des responsabilités, on peut parfaitement concevoir que certains programmes créent leurs propres gisements de données, tandis que d'autres décident de chercher un "lieu sûr" adéquat, exploité et géré par quelqu'un d'autre. Par définition, la décision de gérer des matériaux du patrimoine par l'intermédiaire du gisement de données de quelqu'un d'autre ne dispense pas de la responsabilité finale du programme de préservation concerné.

Ces directives, y compris les notes sur les fonctions ci-après et les notes sur la protection des données du chapitre 16, peuvent servir à définir les critères pour identifier et apprécier des "lieux sûrs" potentiels.

### **8.5.2 L'ingestion**

Le terme d'*ingestion* désigne en général les opérations de réception, de préparation et de transfert des matériaux numériques dans le système d'archivage.

La préparation du matériau afin de l'introduire dans le système d'information d'archivage est une opération d'une importance cruciale pour la manière dont est géré l'ensemble du système. Elle comporte un certain nombre de mesures qui peuvent rendre plus ou moins facile ou difficile la

conservation des ensembles cohérents d'informations archivés dans le système. Ces mesures sont les suivantes :

- mettre en oeuvre des politiques de collecte et des critères de sélection pour évaluer si le matériau est en mesure d'être cherché, ou accepté s'il est présenté ;
- clarifier ou négocier les questions de droits avec les propriétaires de ceux-ci ;
- vérifier la qualité de l'ensemble cohérent d'informations présenté, y compris son intégralité de contenu, la fonctionnalité des parties qui le composent, son authenticité, et la présence éventuelle de matériaux indésirables comme les virus ;
- classer le matériau au moyen d'identifiants uniques ;
- estimer les éléments à conserver et fixer des objectifs de préservation ;
- établir des période de rétention et de révision pour le matériau, en cas de besoin ;
- vérifier, et si nécessaire, enrichir la documentation qui décrit le matériau, y compris les métadonnées techniques et de préservation ;
- évaluer les formats des fichiers et décider s'ils ont besoin d'être changés pour se conformer à la politique du programme en matière de gestion des formats qu'il gèrera (avec ou sans restriction) ;
- si nécessaire, changer les formats de fichier pour se conformer à la politique de formatage ;
- adapter la documentation pour indiquer toutes les modifications.

Une fois que l'objet numérique et ses métadonnées ont été préparés et associés pour constituer un ensemble cohérent d'informations, ils sont sauvegardés dans un système d'archivage.

### **8.5.3 Le système d'archivage**

Un programme de préservation doit fournir un archivage qui conserve, protège et vérifie l'intégrité des ensembles cohérents d'informations stockés, à la fois l'objet numérique et les métadonnées, qu'ils soient stockés comme un train de données unique ou comme trains de données séparés mais reliés.

Pour parvenir à réaliser cela, la fonction d'archivage doit inclure des pratiques qui protègent le train de données contre les modifications involontaires, les détériorations ou la perte. D'où, en général, l'obligation de recopier régulièrement le train de données sur de nouveaux supports et, si nécessaire, de le faire sur des supports d'un type nouveau. D'autres pratiques d'archivage peuvent se révéler nécessaires : vérifier que le train de données n'est pas altéré ; la sécurité du système ; les régimes de sauvegarde qui placent les copies dans des sites éloignés ; et les plans anticatastrophes qui parent à certains aléas comme la perte complète de l'infrastructure de fonctionnement du système.

Bien entendu, il faut des capacités techniques pour mettre sur pied un service d'archivage convenablement sûr et fiable. Un équipement modeste peut suffire à condition que l'équipement et l'ensemble du système soient bien gérés. Plus la quantité, la diversité et la complexité des matériaux sont grandes, plus le système d'archivage doit être élaboré.

#### **8.5.4 La planification de la préservation**

Pour la plupart des matériaux numériques, la préservation de l'accessibilité réclame bien plus que la simple protection des données offerte par la fonction d'archivage. Seul un matériau conservé pour de très courtes périodes pourrait être stocké sans se préoccuper des moyens d'y donner accès.

La fonction de planification de la préservation vise à surveiller les menaces qui pèsent sur l'accessibilité et à préciser l'action à engager pour les prévenir ou y répondre.

Ces menaces sont surtout liées aux changements de technologie qui sous-tendent l'accès. Cette fonction surveille donc ces changements et prend des mesures pour maintenir l'accessibilité malgré ces changements. Fréquemment, l'action impliquera la modification de l'ensemble cohérent d'informations : soit une transformation de l'objet numérique lui-même en un codage différent (comme il arrive dans le passage à un nouveau matériel), soit un changement des métadonnées qui décrivent les moyens d'accès et les liens aux outils d'accès existants.

#### **8.5.5 La gestion des données**

Gérer les ensembles cohérents d'informations archivés engendre ses propres données sur la nature du matériau stocké, sur ce qui peut être accessible et sur la gestion des archives. Il faut gérer ces données pour faciliter l'utilisation des archives et en assurer une administration efficace.

#### **8.5.6 L'accès**

Cette fonction fournit une connexion aux archives, permettant aux utilisateurs de découvrir ce qu'elles contiennent, pour demander le matériau et si nécessaire pour recevoir des copies.

Nombreuses seront les archives dont l'accès sera limité pour certains, voire pour tous les utilisateurs potentiels. Cette fonction d'accès nécessitera des mécanismes de contrôle d'accès.

#### **8.5.7 Liaison et préconisation**

Le programme de préservation doit trouver des moyens de préconiser les bonnes pratiques chez les producteurs, en vue de faciliter la préservation du matériau dont le programme est responsable. Il faut également déterminer quels seront les utilisateurs probables du matériau, de sorte que les dispositifs d'accès puissent être adaptés à leurs besoins et à leurs attentes.

#### **8.5.8 Les fonctions de gestion, d'administration et de soutien**

L'exploitation globale du programme doit être gérée. Cette responsabilité implique d'une part l'élaboration des structures d'une politique et de normes couvrant tout le champ des opérations ; d'autre part la fourniture continue de ressources appropriées et d'une infrastructure qui compte des systèmes techniques adaptés et enfin des méthodes de gestion comme la surveillance et le compte rendu des opérations du programme.

[On trouvera une description plus approfondie et détaillée des responsabilités et des fonctions exposées ci-dessus (sous une forme simplifiée et légèrement modifiée) dans le Reference Model for Open Archival Information Systems (OAIS) (Modèle de référence pour les systèmes d'informations d'archives ouverts), publié en 2002 comme un modèle international d'avant-projet par l'International Standards Organisation (Organisation des normes internationales). Le Modèle de référence de l'OAIS a brillamment réussi à définir à la fois un modèle conceptuel pour gérer les matériaux numériques de valeur durable et un vocabulaire pour en discuter.]

Quiconque envisage de prendre la responsabilité de gérer des matériaux numériques doit comprendre les concepts exposés dans le Modèle de référence lui-même.

Il offre un cadre conceptuel de haut niveau qui peut servir de point de référence à ceux qui conçoivent, utilisent et évaluent des mises en oeuvre réelles. Il faut savoir qu'il ne s'agit pas d'une description détaillée de la mise en oeuvre : il ne fournit pas un ensemble d'instructions sur la manière de préserver l'information numérique. Sa valeur réside dans l'explication de ce que cette préservation exige à un niveau hautement conceptuel, sans se soucier des moyens choisis pour y parvenir.]

## **8.6 Les caractéristiques des programmes de préservation fiables**

La fiabilité et la crédibilité des programmes de préservation numérique sont des points essentiels pour l'ensemble des parties prenantes. Les producteurs, les utilisateurs, les investisseurs et le grand public ont fortement intérêt à veiller à ce que les matériaux du patrimoine numérique soient gérés par des dispositifs auxquels on peut se fier. Ceux qui sont potentiellement responsables des programmes ont également intérêt à évaluer ce qu'ils peuvent offrir et les risques que comporte une prise de responsabilité.

Les programmes de préservation qui offrent une fiabilité sur le long terme doivent présenter les caractéristiques suivantes :

- des responsabilités : un engagement fondamental dans la préservation de matériaux numériques en question ;
- une viabilité organisationnelle, y compris la perspective d'un mandat durable ; un statut juridique digne de celui d'une organisation qui assumerait un rôle de préservation durable ; et une capacité avérée de réunir les ressources, l'infrastructure et les équipes de travail qui puissent gérer la complexité de la préservation numérique ;
- une durabilité financière : l'une des perspectives probables de l'Organisation étant de pouvoir continuer à fournir les ressources nécessaires loin dans l'avenir avec un modèle de gestion durable pour assumer son mandat de préservation numérique ;
- une adéquation entre la technologie et les procédures : l'utilisation de systèmes et de procédures appropriés pour assurer la gestion et la préservation voulues des ressources numériques ;
- un système de sécurité de très haut niveau ;
- une responsabilité procédurale, avec une attribution claire des responsabilités et des mécanismes pour justifier et évaluer l'exécution.

Des dispositifs dotés de ces caractéristiques devraient être fiables. Au fil du temps, le développement de la fiabilité peut être une occasion de tester ces caractéristiques. A long terme, des programmes de certification seront sûrement nécessaires, mais au moment où nous écrivons ces lignes, il n'existe aucun programme de certification pour les dispositifs de préservation numérique. Il revient à ceux qui proposent des programmes de préservation numérique de montrer pourquoi l'on doit se fier à leurs dispositifs, et aux autres parties prenantes d'établir que les dispositifs disponibles sur le marché peuvent offrir un niveau acceptable de fiabilité.

## REFERENCES - pour en savoir plus

### Renvois internes

*L'Ingestion*, voir aussi Prendre le contrôle : chapitre 14.

*L'Archivage*, voir aussi Protéger les données : chapitre 16.

La *Planification de préservation*, voir aussi Maintenir l'accessibilité : chapitre 17.

Les *Programmes exhaustifs et fiables* sont opposés à d'autres possibilités dans Prendre des responsabilités : chapitre 9.

### Références externes

- Consultative Committee for Space Data Systems (2002). Reference Model for an Open Archival Information System.
- (OAIS). CCSDS 650.0-B-1. Blue Book. Issue 1. Janvier 2002. Washington D.C., CCSDS Secretariat, 2002. <http://www.classic.ccsds.org/documents/pdf/CCSDS-650.0-B-1.pdf>
- Research Libraries Group, (2002), Trusted Digital Repositories: Attributes and Responsibilities - An RLG/OCLC Report. Mountain View, Californie, 2002. <http://www.rlg.org/longterm/repositories.pdf>

## Chapitre 9. Prendre des responsabilités

### INTRODUCTION

#### 9.1 Les objectifs

Ce chapitre vise à aider les responsables de programme à établir les responsabilités de préservation qu'ils auront à prendre.

#### 9. En bref

La préservation numérique ne sera effective que si des organisations et des individus acceptent d'en prendre la responsabilité. Prendre des responsabilités signifie mettre des dispositifs en place pour adopter le type de mesures de préservation esquissées dans ces Directives et employer les procédés émergents qui conviennent au fur et à mesure qu'ils deviennent disponibles. Des dispositifs de préservation exhaustifs et entièrement fiables sont nécessaires, mais il arrive qu'ils ne soient pas souvent accessibles au moment voulu ; des programmes sérieux et plus limités peuvent alors apporter une aide précieuse.

### PERSPECTIVE DE GESTION

#### 9.3 Décider de prendre des responsabilités de préservation

Les responsabilités sont un point essentiel dans la préservation du patrimoine numérique. La décision d'assumer ces responsabilités marque le point de départ de l'action.

Etant donné que les coûts qui en résultent pour une organisation peuvent être élevés et que les besoins peuvent se révéler complexes et mal définis, ce n'est pas rien que d'accepter un rôle de responsabilité. Les responsabilités et les fonctions des programmes complets, et les caractéristiques des programmes fiables, telles qu'elles sont décrites dans le chapitre précédent, ne doivent pas être sous-estimées. Elles entraînent un investissement en ressources, en énergie et un projet.

D'autre part, le problème est urgent : de grandes portions du patrimoine numérique seront perdues sous peu à moins que des organisations et des individus n'acceptent d'agir.

La démarche proposée par ces Directives aux organisations intéressées partage leurs décisions de responsabilité en deux groupes de considérations :

- Existe-t-il des bases justifiant de prendre des responsabilités ?
- Si oui, quelle sorte de responsabilités doit-on prendre ?

Dans tous les cas, la qualité des décisions sera influencée par la connaissance approfondie des matériaux considérés, par les tâches à accomplir, les attentes des parties prenantes et les ressources éventuellement disponibles.

### 9.3.1 Première considération : Existe-t-il des bases justifiant de prendre une responsabilité de préservation?

Le tableau 9.1 présente une série de questions qui peuvent servir de point de départ pour l'examen de cette décision.

#### Considération essentielle

Existe-t-il des bases justifiant de prendre une responsabilité de préservation ?

#### Question clé

1. L'activité de l'Organisation implique-t-elle une obligation de préservation, existante ou potentielle, pour de quelconques types de matériaux du patrimoine numérique ? (Demande-t-on à l'Organisation de prendre des responsabilités ?)

#### Questions annexes

(a) Y a-t-il des obligations juridiques existantes ?

#### Comme :

- un dépôt légal ou autre obligation statutaire ;
- des règles d'organisation ;
- des obligations contractuelles.

(b) Les obligations actuelles supposent-elles une extension éventuelle à des matériaux numériques ?

(Par exemple : une bibliothèque de dépôt ne disposant pas d'une juridiction pour les matériaux numériques)

- la responsabilité d'un matériau parallèle ;
- répondre aux besoins d'un client parallèle.

(c) L'Organisation a-t-elle accepté un rôle de conservation pour des matériaux numériques supposant l'attente d'une préservation ultérieure ?

- une donation de données ;
- des données stockées pour des déposants ;
- des données transférées d'un autre programme de préservation.

2. L'Organisation a-t-elle intérêt à prendre des responsabilités de préservation ? (Souhaite-t-elle jouer un rôle ?)

(a) A-t-elle un "intérêt naturel" à identifier les matériaux et à les garder accessibles ?

- pour des utilisateurs ;
- pour de futures recherches ;
- pour une réutilisation ;
- pour la fierté de la communauté ;
- pour la rentabilité.

- (b) A-t-elle un intérêt indirect fondé sur une relation privilégiée avec une communauté particulière ?
- la communauté du producteur ;
  - d'autres personnes concernées.
3. L'Organisation a-t-elle, ou peut-elle, acquérir les aptitudes nécessaires pour prendre une responsabilité de préservation ?
- (a) Peut-elle apporter ce qui sera nécessaire pour exercer une responsabilité ?
- un engagement et un projet ;
  - les ressources ;
  - les connaissances et les compétences ;
  - les contacts ;
  - la crédibilité ;
4. La responsabilité appartient-elle véritablement à quelqu'un d'autre ?
- (a) Y a-t-il quelqu'un d'autre qui exerce déjà cette responsabilité, ou qui pourrait le faire ?
- Quelqu'un qui le fait déjà ;
  - Quelqu'un à qui l'on a déjà demandé de le faire ;
  - Quelqu'un qui a un "intérêt naturel" à le faire ;
  - Qui a les aptitudes pour le faire.

**Tableau 9.1** *Considérations à examiner pour déterminer s'il existe bien une responsabilité de préservation*

**9.3.2 Seconde considération : Quelle sorte de responsabilité doit-on prendre ?**

Ces Directives recommandent vivement aux programmes de préservation de s'efforcer d'être conformes aux critères d'exhaustivité et de fiabilité décrits dans le chapitre 8. Ce sont des références importantes pour tous les programmes.

Toutefois, de nombreuses organisations peuvent décider qu'elles ont une responsabilité de préservation mais découvrir qu'elles ne sont pas en mesure de remplir ces critères et se demander s'il y a une place pour des programmes de préservation plus modestes. N'y a-t-il d'autre alternative que l'action pour ceux qui peuvent s'y conformer et l'inaction pour ceux qui n'en sont pas capables ?

Dans de nombreux environnements, il peut n'y avoir personne à même d'assumer une responsabilité de préservation entière et fiable. La seule chance de survivance d'un patrimoine numérique dépendra alors de quelqu'un qui entreprend une action limitée, incertaine mais avertie, tant que cela est possible. Cela permet au moins de gagner suffisamment de temps pour que soient mis en place des dispositifs plus fiables.

Même là où des programmes complets et fiables sont disponibles, un rôle important échoit à des programmes qui assurent la responsabilité de certains processus seulement, tout incapables qu'ils soient de prendre la responsabilité des processus dans leur intégralité. En fait, la plupart des programmes de préservation étendus n'ont une chance de durer que s'ils trouvent des partenaires désireux et capables de jouer un rôle limité mais complémentaire.

Outre qu'ils offrent divers degrés d'exhaustivité et de fiabilité, les programmes de préservation se distinguent par la gamme des matériaux qu'ils s'efforcent de préserver et par la durée pendant laquelle s'exerce leur responsabilité. Les programmes qui assurent une préservation exhaustive et fiable pour des gammes de matériaux très restreintes et des périodes de temps très limitées ont indéniablement un rôle précis à jouer.

C'est là un point extrêmement important. Il l'est d'autant plus que l'affirmation souvent réitérée que la préservation numérique passe par un engagement à très long terme risque de faire obstacle à la préservation en décourageant les organismes qui sont bien placés pour mener une action sur le court terme lorsqu'elle est nécessaire. Moyennant une bonne planification de leurs relais, les organisations à même de jouer un rôle efficace mais limité dans le temps peuvent aider le nombre plus restreint d'organismes aptes à s'engager dans une action de conservation soutenue véritablement de longue durée.

Le tableau 9.2 propose une gamme de niveaux de responsabilité en regard de quatre continuums clés : l'éventail des matériaux préservés ; la période de temps pendant laquelle seront assumées les responsabilités ; l'étendue des fonctions et des responsabilités essentielles qui seront assumées ; et l'existence de caractéristiques de fiabilité.

Il se peut que les communautés concernées par la préservation numérique aient besoin d'élaborer leurs niveaux de responsabilité et leurs critères propres pour les programmes chargés de leur patrimoine.

Il se peut également que les organisations qui prennent des responsabilités de préservation, et leurs partenaires, trouvent utile d'établir le niveau de leur prise de responsabilité par rapport à ces continuums.

Toutes les propositions pour prendre des responsabilités de préservation doivent être bien documentées, mûrement réfléchies et fondées sur une vision claire de ce qui doit être accompli - même si la manière dont on va surmonter tous les obstacles et relever tous les défis n'est pas encore tout à fait claire.

#### 1. *L'éventail des matériaux*

##### **Programme restreint**

très restreint.

##### **Programme sélectif**

à large couverture, collecté de manière exhaustive.

## **Programme étendu**

### 2. *La période de temps*

#### **Programme initial**

seulement jusqu'à ce que soit modifiée la technologie.

#### **Programme de conservation**

seulement jusqu'à l'arrêt de l'utilisation ;

pour un nombre limité d'années ;

"à perpétuité".

#### **Programme sur le long terme**

### 3. *L'étendue des fonctions et des responsabilités*

#### **Programme partiel, non complet**

fonctions restreintes.

#### **Programme complet**

fonctions complètes.

### 4. *Le niveau de fiabilité*

#### **Programme non fiable**

caractéristiques de fiabilité limitées

#### **Programme entièrement fiable**

toutes les caractéristiques de fiabilité.

### **Tableau 9.2** *Niveaux de responsabilités - quelques continuums possibles*

## **9.4 Planification de la préservation sur le long terme**

Tous les programmes, mais en particulier ceux qui sont incapables de proposer des engagements sur le long terme, doivent s'efforcer de mettre en place une sorte de *mécanisme à sécurité intégrée*. L'objet de ce dispositif est de fournir une bonne prospective de préservation qui se poursuive au-delà de leur propre engagement, si cela devenait nécessaire.

Des dispositifs à sécurité intégrée s'étendent sur un continuum, depuis l'engagement de trouver quelqu'un d'autre pour prendre la relève avant de mettre au rebut les matériaux numériques jusqu'aux dispositions mandatées juridiquement qui permettent à un organisme de reprendre la gestion des données détenues par un autre organisme si celui-ci ne parvient pas à assumer ses responsabilités de préservation.

La responsabilité d'instaurer et d'entretenir des plans de succession incombe évidemment à celui qui gère le matériau, mais il peut y avoir d'autres acteurs qui acceptent certaines responsabilités. Les organismes qui croient qu'on leur demandera de prendre en charge la tâche de préservation et ceux qui souhaitent voir les matériaux accessibles sont également concernés et peuvent avoir à jouer un rôle actif en engageant des négociations avec les conservateurs actuels.

## 9.5 Quelques principes pragmatiques de responsabilité

Face aux défis impressionnants de la responsabilité de préservation, il peut être utile de prendre en considération certains principes pragmatiques :

- quelqu'un doit prendre des responsabilités : si personne ne le fait les chances de survivance de tous les matériaux particuliers sont très minces ;
- un individu n'a pas à tout faire : les responsabilités peuvent être partagées. Comme le montrent ces Directives, il y a plus de responsabilités qu'il n'en faut pour un seul programme de préservation. De nombreuses tâches, comme celle de décider ce qui doit être préservé, sont mieux gérées en association avec d'autres. S'il n'y a personne à même de partager les responsabilités, les programmes de préservation doivent évaluer avec réalisme les responsabilités qu'ils peuvent assumer seuls ;
- il ne faut pas tout faire sur-le-champ : développer toutes les composantes d'un programme de préservation complet de grande envergure prend du temps. Il est bon d'aborder la tâche avec un sentiment d'urgence, mais il peut aussi être nécessaire de chercher des moyens de gagner du temps. Il conviendrait donc de donner la priorité aux problèmes qui doivent être traités, ou au matériau qui nécessite de l'attention. Cette démarche peut impliquer la recherche de matériaux facilement gérables, "à portée de la main". Il faut s'attaquer à certains problèmes sans tarder ; d'autres seront résolus par étapes ; et d'autres encore peuvent attendre ;
- la prise de responsabilité ne saurait être perpétuelle ; il y a place, encore une fois, pour les contributions d'une durée limitée à un programme de préservation globale, dès lors que les limites dans le temps sont formellement définies ;
- une responsabilité limitée ne signifie pas pour autant qu'elle est néfaste : les programmes de préservation nécessitent parfois de prendre des mesures avant que tous les problèmes soient résolus et toutes les techniques mises en place, mais ils doivent également tenter de minimiser ce qui risque de rendre les efforts de préservation ultérieurs plus difficiles ;
- quelqu'un doit tenir un rôle directeur : même lorsque les responsabilités sont partagées, les progrès dépendent généralement du fait qu'au moins un des partenaires accepte d'être à la tête des responsabilités.

## 9.6 Qui pourrait prendre la responsabilité

Qui pourrait prendre la responsabilité d'établir et de gérer des programmes de préservation pour les matériaux du patrimoine numérique ? Entre autres éventualités, on peut envisager d'étendre le rôle des institutions de la "mémoire" comme les bibliothèques, les archives et les musées ; ou bien d'instaurer une nouvelle sorte d'institution axée exclusivement sur la préservation des matériaux numériques ; d'ajouter un rôle de préservation à une gamme d'autres "gardiens" potentiels déjà engagés dans la gestion des matériaux numériques ; ou enfin une combinaison des trois.

### **9.6.1 Le rôle des institutions établies du patrimoine**

Lorsqu'on a commencé à discuter de la manière dont il faudrait préserver les matériaux numériques, on a souvent prétendu que la technologie numérique bouleversait le paysage : les institutions existantes découvrirait qu'elles n'ont aucun rôle à jouer dans la gestion de ces matériaux-là.

Il est encore beaucoup trop tôt pour juger si la démission "annoncée" des institutions de l'information et culturelles traditionnelles aura lieu, mais l'expérience montre que l'annonce nécrologique est prématurée. Lorsque l'on cherche des organisations susceptibles d'offrir ce qui est nécessaire, les institutions qui gèrent déjà des matériaux du patrimoine non numérique semblent présenter de nombreux avantages. Nombre d'entre elles proposent :

- une compétence pour identifier les matériaux du patrimoine importants ;
- l'habitude de travailler avec les communautés d'utilisateurs ;
- l'habitude de travailler avec les propriétaires de droits ;
- une compétence et des réseaux informatiques spécialisés dans l'organisation et la description des matériaux du patrimoine de sorte qu'ils puissent être trouvés et compris ;
- un engagement pour leur préservation sur le long terme ;
- au moins quelques compétences et une infrastructure pertinentes qui pourraient être exploitées pour encourager la gestion du bien numérique et au moins ;
- une perspective de mandat durable de la part de leurs groupes pour la gestion et la préservation du patrimoine numérique.

A ces bases prometteuses, certaines institutions ont su adjoindre un rôle directeur dans la recherche de moyens pratiques pour préserver le patrimoine numérique.

Cela ne signifie pas nécessairement que toutes les institutions qui ont un rôle patrimonial traditionnel doivent s'efforcer de devenir les gestionnaires d'un patrimoine numérique. Dans certains cas, les ressources et les compétences ne sont tout simplement pas disponibles ; tandis que dans d'autres, le rôle existant de ces institutions est si important et si accaparant qu'elles ne vont pas sacrifier ce dont elles ont déjà la garde au profit de ce qui peut être un matériau numérique beaucoup moins important.

Cela ne signifie pas non plus que les institutions patrimoniales existantes sont les seules organisations qui ont besoin de gérer les matériaux du patrimoine numérique.

Mais cela laisse entendre que les institutions patrimoniales existantes sont de bons points de cristallisation autour desquels peuvent se développer des programmes de préservation du patrimoine numérique. De telles institutions doivent tenir compte des forces existantes qu'elles apportent à la gestion des matériaux du patrimoine numérique, souvent en partenariat avec d'autres qui offrent une palette de compétences et de points de vue nouveaux.

Parmi les institutions patrimoniales existantes, les bibliothèques nationales, les organismes nationaux d'archivage et d'autres institutions maîtresses dans divers secteurs peuvent jouer un rôle particulièrement important en lançant des programmes de préservation. Ce phénomène est déjà apparu dans de nombreux pays.

### ***9.6.2 Les rôles de nouveaux types d'organismes de préservation numérique***

Pour certains, de nouvelles institutions seront nécessaires pour prendre en charge la préservation du patrimoine numérique. Ces organismes offriraient probablement les compétences de spécialistes et des équipements consacrés aux matériaux numériques, voire à la préservation, plutôt qu'un éventail plus large de fonctions que les institutions de collecte existantes accomplissent, telle la mise en ordre et l'interprétation des matériaux, et la promotion de leur utilisation.

De nombreuses archives de données correspondent déjà à ce schéma, puisqu'elles existent uniquement pour gérer et préserver les matériaux numériques. Elles ont souvent l'avantage paradoxal de se borner à un registre limité de matériaux et de tâches de gestion, tout en offrant des services à un éventail plus large de communautés productrices de données.

### ***9.6.3 Les rôles d'autres gardiens fiables***

Qui d'autre pourrait assumer le rôle de gardiens fiables ? De même qu'il est trop tôt pour évaluer le rôle sur le long terme des bibliothèques, des archives et des musées, il est prématuré d'essayer de dresser une liste définitive des autres acteurs qui pourraient jouer un rôle important. Toutefois, certaines possibilités s'imposent déjà :

- les universités et autres instituts de recherche et d'études ont logiquement intérêt à assurer un accès durable à certains types de matériaux numériques ; de tels établissements peuvent avoir une viabilité sur le long terme et l'infrastructure technique pour jouer un rôle de préservation ;
- les éditeurs et les créateurs du contenu numérique s'intéressent pour toute une série de raisons à la gestion et à l'accessibilité durables. Dans de nombreux cas, cet intérêt dépasse largement les considérations commerciales immédiates pour un investissement sur un plus long terme dans l'échange d'idées et le capital intellectuel et culturel qu'encourage l'accessibilité durable. Certains éditeurs et créateurs peuvent être désireux et à même de fournir l'infrastructure nécessaire à la conservation des matériaux du patrimoine numérique auxquels ils sont attentifs.

## **9.7 Affirmer ses responsabilités**

Une fois que les programmes de préservation ont établi le type de responsabilités qu'ils prendront, il est capital qu'ils énoncent leurs intentions. Cela facilite le travail avec les autres, réduit l'éventualité que les mêmes efforts soient inutilement faits par d'autres et donne une vue plus claire des matériaux susceptibles de survivre ou non.

Cette déclaration explicite de responsabilités se doit d'être également réaliste : des prétentions excessivement optimistes laisseraient supposer un niveau de sécurité de préservation qui n'existe pas et les autres programmes risqueraient de n'être pas en mesure d'intervenir au dernier moment pour sauvegarder un matériau qu'ils croyaient relever de la responsabilité de quelqu'un d'autre.

## REFERENCES - **pour en savoir plus**

### **Renvois**

*Les programmes de préservation complets et fiables*, voir aussi Comprendre les programmes de préservation numérique : chapitre 8.

*Les problèmes de droits dans la prise de responsabilités*, voir aussi Gérer les droits : chapitre 15.

## Chapitre 10. Gérer les programmes de préservation numérique

### INTRODUCTION

#### 10.1 Les objectifs

Le propos de ce chapitre est de proposer quelques domaines clés dans les précautions à prendre pour bien gérer les programmes de préservation. Les responsables expérimentés sont déjà familiarisés avec la majeure partie de cet exposé.

#### 10.2 En bref

Les programmes de préservation réclament une bonne gestion : elle s'appuie souvent sur des compétences de gestion générale, comme le façonnage de programmes adaptés aux priorités et au contexte d'un cas particulier, et sur la capacité à prendre les décisions requises au moment voulu. Les programmes de préservation numérique présentent quelques problèmes de gestion spécifiques liés à leur caractère évolutif, à l'éventail des parties prenantes, et aux conséquences sur le long terme de décisions prises aujourd'hui.

### PERSPECTIVES DE GESTION

#### 10.3 La nécessité de gérer

Comme tout programme, les programmes de préservation numérique réclament une gestion cohérente, qui n'a rien d'un art mystérieux : elle repose sur de bonnes compétences de gestion générale, alliées à une connaissance du sujet et une compréhension des questions techniques suffisantes pour envisager des possibilités réalistes et prendre les bonnes décisions. L'exposé qui suit est centré sur les problèmes qui sont spécifiquement du ressort de la préservation numérique.

#### 10.4 Ce dont la gestion du programme doit traiter

##### 10.4.1 Les décisions

Une bonne gestion se résume souvent à savoir quelles sont les décisions à prendre et à les prendre au moment requis. Les décisions importantes n'ont pas besoin d'être prises sur-le-champ, mais en fin de compte elles se révéleront indispensables sur les points suivants :

- s'il faut ou non s'impliquer dans la préservation des matériaux numériques ;
- la mission du programme ;
- la portée de l'implication : le type de matériau à inclure, la dimension à donner au programme : s'il vise à être exhaustif et fiable ou se veut plus modeste ;
- où se procurer les services dont le programme aura besoin ;
- choisir les ressources que l'on rendra disponibles et savoir comment le programme sera maintenu ;
- quelles sont les structures d'organisation nécessaires pour le soutien du programme ;
- avec qui le programme travaillera-t-il ;
- quelles sont les questions auxquelles il faut s'attacher en priorité ;
- comment l'accessibilité sera-t-elle maintenue ;

- quels systèmes de relais à mettre en place comme mécanisme à sécurité intégrée.

#### **10.4.2 Les risques et la gestion du risque**

Les programmes de préservation doivent chercher à comprendre et à faire face aux menaces qui mettraient en péril l'accessibilité continue et les autres aspects de leur mission. Une méthode de gestion du risque fournit la base voulue pour déterminer les risques à surveiller et planifier une action qui abaissera le niveau du risque.

Il existe de nombreux modèles de gestion du risque qui conviennent. Nous en proposons un, assez simple mais efficace, dans le tableau 10.1.

L'analyse du risque, même si elle est menée de façon informelle, est utile à bien des égards. Elle aide à :

- reconnaître :
  - les menaces les plus urgentes (comme la disparition des publications Web ; les défaillances des supports magnétiques ; le remplacement imminent de l'équipement ou du logiciel ; un changement dans les dispositifs d'un organisme gouvernemental qui menacera les systèmes de conservation des archives) ;
  - les menaces qui peuvent ne pas nécessiter une action immédiate (comme la tombée en désuétude d'un format de fichier omniprésent et normalisé comme le TIFF, dont les conséquences doivent être gérables lorsqu'apparaît une norme de remplacement) ;
  - les menaces sur lesquelles le programme peut n'avoir aucune prise (comme les impératifs commerciaux des producteurs) ;
  - les menaces qui sont si urgentes mais si imparables que le programme peut décider de dégager ses responsabilités en la matière (comme le refus de la part de propriétaires de droits de permettre toute forme d'accès ou de copie de préservation à l'avenir) ;
- déterminer :
  - où affecter les ressources ;
  - quelles mesures prendre en priorité ;
  - quand une action peut être nécessaire ;
  - quelle action de soutien est nécessaire pour traiter les risques prioritaires ;
- planifier à l'avance ;
- justifier les décisions ;

L'évaluation des risques est particulièrement utile si, voyant plus loin que les risques manifestes, elle englobe :

- les risques liés à l'action proposée pour venir à bout de la menace.

Par exemple, il se peut que le programme n'ait pas les compétences, les ressources, les permissions nécessaires, etc., d'où la nécessité d'envisager une action préalable pour traiter la menace prioritaire.

Les causes de la menace originelle. Par exemple, des éditeurs Web ne sont pas forcément au courant des mesures qu'ils peuvent prendre ; des producteurs peuvent mal utiliser les normes ; ou bien se trouver au bord de la faillite commerciale. Cette analyse pourrait conduire à une action qui traite les causes, comme des campagnes d'éducation, l'élaboration de normes de concert avec les producteurs, ou l'élaboration d'indicateurs de faillite commerciale imminente comportant des signes qui signalent notamment que des sites Web ne sont pas maintenus ou que des projets arrivent en bout de course

### **Les démarches**

1. *L'identification du bien* : identifier ce qui doit être protégé, aussi précisément que possible.
2. *L'identification de la menace* : identifier les menaces qui semblent présenter un risque pour des objectifs du programme.
3. *L'évaluation de la probabilité* : estimer la probabilité de la concrétisation de chaque menace.
4. *L'évaluation des conséquences* : estimer les conséquences probables si la menace se concrétise.
5. *L'évaluation du niveau de risque* : calculer le niveau du risque en conjuguant la probabilité et les conséquences.
6. *La mitigation* : proposer une action qui pourrait réduire la probabilité ou les conséquences de la menace, ou les deux.
7. *Le seuil de risque* : déterminer si le niveau de risque est acceptable avec ou sans action de mitigation.
8. *L'attribution de la propriété* : déterminer qui a la responsabilité d'agir et toutes les contraintes.
9. *L'établissement des priorités* : comparer les niveaux de risque pour les menaces identifiées et décider les risques auxquels on doit donner la priorité.
10. *Le contrôle de la réalité* : déterminer si le risque et les évaluations de priorité s'accordent assez bien avec les attentes.
11. *Les déclencheurs de l'action* : déterminer si l'action est immédiatement nécessaire ; sinon, identifier certains signes qui indiqueront le moment où l'action est nécessaire.

### **Exemple concret de travail**

Publication en ligne stockée sur un site Web géré par quelqu'un d'autre.

L'accès à une version particulière de la publication en ligne sera perdu parce que le propriétaire écrase les anciennes versions avec de nouvelles versions.

Une concrétisation très probable de la menace, fondée sur les antécédents du site.

Aboutissant probablement à la perte totale de l'ancienne version, puisque le propriétaire ne semble pas conserver d'archives des versions écrasées.

Risque élevé - la menace se concrétisera probablement et aboutira à une perte complète.

Les options - joindre le propriétaire du site et lui proposer de faire des copies archivées ; négocier l'autorisation de faire immédiatement des copies ; ou faire des copies avant que les versions ne soient écrasées.

Le matériau est jugé important de sorte que le niveau de risque justifie de prendre des mesures de mitigation.

Le propriétaire peut être responsable mais il n'est pas forcément disposé à agir ; le programme peut assumer la responsabilité, mais il aura sans doute besoin d'une autorisation

Une priorité essentielle par rapport aux autres risques.

Aucun doute ne subsiste - l'analyse est "significative".

Le propriétaire est d'accord pour que le programme de préservation réalise la reproduction immédiatement avant que chaque version soit écrasée, mais il ne peut pas fournir un calendrier. Le programme décide de joindre régulièrement le propriétaire pour se renseigner sur le calendrier des mises à jour, et évaluera si ce moyen se révèle être un indicateur suffisant.

*Tableau 10.1. Un modèle de gestion du risque simple*

#### **10.4.3 Les relations entre parties prenantes**

Un large registre de personnes sont directement concernées par les matériaux du patrimoine numérique et la manière dont ils sont gérés. Figurent parmi eux ceux qui créent ou diffusent les matériaux, ceux qui ont besoin de les utiliser aujourd'hui ou dans le futur, et ceux qui mentionnent les matériaux, en supposant qu'ils resteront disponibles. Sans forcément apparaître au premier plan, certains d'entre eux jouent un rôle déterminant dans les coulisses, tels les fabricants de matériel et de logiciels, les bailleurs de fonds et les organismes autorisés à contrôler l'accès à certains matériaux. Certains intéressent au premier chef le programme de préservation lui-même, en tant que partenaires et collaborateurs éventuels, les organismes chargés des normes et les chercheurs qui élaborent de nouvelles méthodes de préservation numérique.

Les programmes de préservation doivent identifier ceux que leurs objectifs intéressent, ou qui peuvent influencer sur eux. L'évaluation du risque est susceptible d'identifier les personnes concernées les plus importantes et le type de relation que le programme a besoin de développer avec elles. (Certains points qui s'y rapportent sont abordés dans les chapitres suivants relatifs aux producteurs, aux propriétaires de droits et à d'autres programmes de préservation.)

Les ouvertures favorables à la préservation sont beaucoup plus limitées que pour le patrimoine non numérique. Aussi, les programmes de préservation doivent-ils se montrer beaucoup plus entreprenants dans la recherche de relations avec les producteurs des matériaux du patrimoine. Ils peuvent aussi s'efforcer de trouver les moyens d'influencer les parties prenantes "occultes" dont les décisions peuvent avoir des conséquences cruciales.

#### **10.4.4 La durabilité et les modèles d'entreprise**

Les programmes de préservation fiables doivent être maintenus sur de longues périodes, aussi réclament-ils des modèles d'entreprise qui garantissent que les ressources nécessaires continueront d'être disponibles. Malheureusement, de telles garanties sont rares dans le monde réel. La majeure partie des programmes doit survivre avec moins de certitude.

Surtout au stade initial, la durabilité présente un dilemme : les programmes doivent finir par être durables, mais souvent ils ignorent encore les ressources qui seront nécessaires pour assumer leurs responsabilités, et celles qui seront disponibles une fois qu'un programme réussi a vu le jour. Les responsables de programmes ont ce défi à relever : trouver des modèles d'entreprise sur le long terme, mais aussi des arrangements de financement sur le court terme, et identifier le bon moment pour passer de l'un à l'autre. (En fait, certains responsables découvriront qu'ils ne peuvent construire un programme à long terme qu'en s'appuyant sur une succession d'arrangements financiers à court terme.)

Certains des modèles d'entreprise ordinairement envisagés pour les programmes de préservation comportent :

- un financement de l'ensemble de la communauté, souvent par l'impôt ou l'attribution de subventions provenant de sources particulières comme les fonds de loterie ;
- un financement par une communauté spécifique concernée par le programme, comme le projet d'une communauté locale, ou une souscription auprès d'instances professionnelles ;
- un financement central par une organisation mère dans le cadre des frais de fonctionnement, telle qu'une bibliothèque d'université ou des archives commerciales ;
- des paiements du matériau par ses utilisateurs ;
- des paiements par les producteurs qui déposent des matériaux avec le programme ;
- un parrainage ou une aide philanthropique ;
- un subventionnement transversal par d'autres activités au sein d'une organisation ;
- une conjugaison de méthodes.

Chacun de ces modèles est plus ou moins approprié ou praticable pour tel ou tel type de programmes. On peut également créer d'autres modèles pour parvenir à la durabilité.

## **10.5 Ce dont les responsables de programme ont besoin pour accomplir leur travail**

### ***10.5.1 Des informations***

Les normes et les techniques de la préservation numérique évoluent rapidement (quoique peut-être pas aussi vite qu'on en aurait besoin), et les responsables des programmes doivent trouver les moyens de garder leurs connaissances à jour.

Il existe un certain nombre d'initiatives internationales, conçues précisément dans ce sens, notamment les revues spécialisées, les sites Web et les moyens d'accès au sujet figurant dans la liste bibliographique. Malgré quelques chevauchements, ces initiatives se complètent les unes les autres : elles offrent d'excellents moyens de rester au courant des évolutions.

Il semble qu'il y ait aussi une occasion d'explorer les relations de solidarité entre des programmes de différents niveaux d'expérience et de compétences. La plupart des programmes existants sont prêts à partager des informations et des idées mais se retrouvent souvent submergés par les demandes. Des arrangements de partage d'informations formels peuvent alléger cette charge tout en offrant des avantages réels à ceux qui élaborent de nouveaux programmes.

### ***10.5.2 Le soutien d'entreprise***

Les programmes de préservation numérique commencent souvent sous la forme d'expériences et de projets dont la mission est conjecturale et l'avenir incertain. Ils peuvent facilement être écartés lorsqu'ils apparaissent comme des compléments externes importants mais gênants d'organisations déjà trop étendues. Il faut souvent un certain temps pour que le déroulement d'un ensemble bien défini d'opérations émerge et que le programme prenne tournure.

Les jeunes programmes ont particulièrement besoin de mentors "maison" ou de sponsors pour défendre leur cause et parler en leur faveur dans les forums d'entreprise. Ils réclament également des moyens de se connecter avec d'autres dans l'Organisation, tels que les comités de gestion, qui peuvent garder le programme aligné avec les directions d'entreprise, tout en fournissant un retour d'informations sur l'évolution et les problèmes. Ce soutien d'entreprise doit tenir compte du fait que les programmes de préservation exigent le plus souvent une grande quantité de ressources et entraînent des problèmes techniques et organisationnels complexes.

### ***10.5.3 Les coûts des ressources***

La disponibilité des ressources est toujours une contrainte cruciale. Il est essentiel de plier les ambitions du programme à une idée assez réaliste de ce qu'il est raisonnablement possible d'atteindre : tout faire est impossible !

Il est difficile d'évaluer les coûts sur le long terme de la préservation numérique. Même si cela apparaît peu satisfaisant du point de vue de la planification, il peut être nécessaire d'estimer les coûts attendus sur une période allant du court au moyen terme, cinq ans par exemple, et d'utiliser les connaissances acquises dans ce laps de temps pour disposer d'estimations plus précises des coûts sur le long terme. Toutefois, il risque d'être toujours difficile d'estimer les coûts d'actions à longue échéance qui sont encore incertains, tels les coûts de transfert.

Il devient nettement plus facile de prévoir certains coûts après une expérience de quelques années dans la collecte du matériau, sa préparation pour l'archivage et sa protection.

On peut raisonnablement émettre certaines hypothèses de coûts :

- les coûts de développement sont susceptibles d'être élevés, en fonction de la hauteur des ambitions du programme depuis le départ. La conception des systèmes est un investissement nécessaire à la longue, mais il entraîne souvent un coût d'installation important ;
- il y a évidemment les dépenses courantes liées au personnel, aux locaux, à l'approvisionnement en énergie, à l'utilisation du réseau, aux télécommunications, et aux supports d'archivage comme les disques, les bandes et les consommables. Bien que souvent financés comme des dépenses d'investissement, l'équipement et les logiciels sont à considérer comme des dépenses courantes car il faut les renouveler régulièrement ;
- les coûts de personnel liés au travail avec les producteurs peuvent être élevés car il faut traiter de nouveaux problèmes à chaque modification technologique. Quant aux coûts des droits de négociation, ils varient selon le degré de complexité de la propriété de ces droits et selon que l'on doit ou non les acquérir ;
- les coûts de l'identification et de la sélection de matériaux pour la préservation peuvent se révéler faibles à l'unité, mais il arrive que les unités soient nombreuses. Une approche non

sélective peut réduire les coûts, mais elle s'ajoute aux coûts globaux d'une préservation sur le long terme, car il faut alors stocker, traiter, préserver et rendre accessibles un plus grand nombre de matériaux. En ce qui concerne la sélection, les interventions humaines sont coûteuses là où les frais de main-d'oeuvre sont élevés ; la prise de décisions automatisée peut réduire les coûts s'il est possible de répartir les coûts d'installation élevés sur une grande quantité de matériaux, et si l'automatisation de ce qui sont souvent des appréciations humaines complexes se révèle réalisable ;

- les coûts de la collecte et du transfert des matériaux peuvent être bas à l'unité, mais des programmes vastes risquent d'entraîner des coûts de transfert globaux importants, notamment des frais de transmission élevés là où la collecte automatisée passe au crible et télédécharge de grandes quantités de données. La vérification du contrôle de la qualité risque aussi de coûter cher, sauf si l'on parvient à l'automatiser ;
- convertir le matériau en une gamme limitée de formats normalisés peut être bon marché si la conversion est facile, mais coûteux si une intervention manuelle ou une correction individuelle se révèlent nécessaires. Toutefois, il y aura des gains considérables à longue échéance si l'on se contente d'une gamme limitée de formats ;
- les coûts de description du matériau et du rajout de métadonnées sont susceptibles d'être élevés en raison du volume d'informations à enregistrer et de la difficulté à les trouver dans certains cas. Les producteurs pourraient les réduire sensiblement s'ils utilisaient eux-mêmes des structures plus normalisées et créaient des métadonnées et une documentation de qualité. Pour des formats très normalisés comme ceux d'un usage répandu pour les versions d'archivage d'images et d'audio, on réussirait à réduire les coûts par la collecte automatisée des métadonnées depuis les fichiers et pendant tous les processus de production ;
- les coûts de stockage sont théoriquement bas et dégressifs, mais au total ils reflètent le volume de données à archiver, souvent important. Dans l'estimation du stockage et des coûts de traitement, il faut prendre en compte les copies de secours et les multiples versions du matériau. Le coût d'une copie régulière des données sur de nouveaux supports est généralement ristourné par des augmentations régulières dans les densités d'archivage ;
- les coûts de fourniture des moyens d'accès - comme l'analyse des structures de données, l'écriture d'un nouveau code pour le transfert ou la recherche d'émulateurs, le contrôle de la qualité, etc., - seront élevés ou faibles selon le degré de difficulté que présente le processus. Par exemple, si l'on parvient à élaborer un outil logiciel de conversion qui fonctionne automatiquement pour des millions de fichiers similaires, les coûts unitaires seront extrêmement bas, même si le coût total du temps informatique reste encore important. En revanche, les stratégies qui marchent élément par élément seront très coûteuses ;
- on ne saurait ignorer les coûts informatiques pour stocker et fournir d'importantes quantités de copies de consultation, ou pour les créer sur demande. Ils ne devraient pas entrer en général dans les coûts de préservation, mais ils peuvent être pris en charge par les programmes de préservation dans le cadre de leur fonction commerciale.

## **10.5.4 Les problèmes de ressources**

### **10.5.4.1 Le personnel à recruter**

La quantité de personnel nécessaire dépendra des dimensions et du type de programme. Là où le matériau est très similaire et bien organisé, et où les tâches de préservation sont simples, il est possible d'automatiser la plupart des processus et de réduire les besoins de personnel au strict minimum. Par contre, là où une intervention humaine répétée se révèle nécessaire, le bon fonctionnement du programme exigera un personnel dûment équipé. Il est très malaisé d'automatiser des processus qui réclament une appréciation subjective, quoique cela ne soit pas impossible. Même des programmes très automatisés, cependant, finiront par avoir besoin de personnel pour faire face aux modifications qui surviennent dans l'environnement d'exploitation et à la nécessité de remplacer les systèmes.

Trouver un personnel doté des compétences requises tient souvent du défi. Il existe quelques programmes de formation pour les "conservateurs numériques", voire pour les responsables de collection numérique. Or, la préservation numérique fait appel à diverses compétences déjà existantes : la technologie de l'information, la gestion de la préservation, la gestion de la collecte et la conservation de l'information. Il est fort possible de trouver du personnel qui soit compétent dans un ou même plusieurs de ces domaines et qui ait les capacités autant que le désir de développer ses compétences. Autre possibilité : compléter les compétences internes dont on dispose avec d'autres qui sont disponibles chez les prestataires de service appropriés.

Les responsables de programme ont besoin d'avoir :

- de bonnes compétences pour résoudre les problèmes et la capacité de régler des problèmes complexes pour lesquels on ne trouve pas actuellement de solutions sur le long terme ;
- une démarche dynamique qui tient compte des problèmes sur le court, le moyen et le terme ;
- une perception suffisante des aspects techniques, de préservation, d'entreprise, juridiques et politiques pertinents ;
- une capacité de réflexion critique, mais aussi d'ouverture aux nouvelles idées et d'adaptation au changement ;
- des capacités de collaboration ainsi que le souci de partager l'information et de rechercher des moyens de travailler avec autrui.

### **10.5.4.2 L'équipement**

Les programmes de préservation demandent des systèmes et des outils perfectionnés. Toutefois, on peut obtenir, avec un équipement informatique très modeste et déjà disponible, des petits programmes aux processus de fonctionnement très peu automatisés et aux supports que l'on peut entreposer sur des étagères. Là où les services sont sous-traités par un fournisseur extérieur, le programme lui-même n'a pas forcément à fournir une grande partie de l'infrastructure technologique.

Les systèmes et les outils susceptibles d'être nécessaires comprennent ceux pour :

- le stockage et la gestion du matériau de la collection ;
- le stockage et la gestion des métadonnées ;
- la gestion de la transmission du matériau à la collection, incluant éventuellement la collecte ;
- la gestion des informations au sujet des droits et la gestion de l'accès conformément aux droits ;
- le stockage et la gestion des outils d'accessibilité comme les logiciels d'origine, les modules d'extension, les programmes de conversion et d'émulation ;
- les recherches faites par les utilisateurs ;
- la réalisation de copies adéquates disponibles pour les utilisateurs ;
- la gestion des nombreux processus décrits dans ces directives, en particulier ceux qui fournissent de l'information sur la gestion, copient le matériau d'un endroit ou d'un format à un autre, ou réclament un contrôle du processus automatisé.

L'acquisition de systèmes adéquats est une tâche importante, qui demande d'être particulièrement attentif aux spécifications et à l'évaluation des options. Certains programmes existants sont prêts à partager des informations sur les spécifications qu'ils utilisent et leur expérience dans l'acquisition de systèmes adaptés.

## **10.6 Les outils utiles pour les programmes de préservation**

### ***10.6.1 Les normes***

Les normes sont présentes derrière presque toutes les activités des ordinateurs, ainsi constituent-elles une base essentielle pour la création et l'utilisation des matériaux numériques. Toutefois, elles n'ont pas été jusqu'alors de grands catalyseurs de la préservation. Or, si elles sont bien choisies et bien utilisées, elles peuvent apporter des contributions importantes dans plusieurs domaines :

- la création de matériaux numériques qui doivent être assez simples à préserver. Certains formats de fichiers (eux-mêmes normalisés) se sont révélés si largement utiles qu'ils ont été adoptés par des groupes de créateurs comme étant d'un meilleur usage, avec l'espoir qu'on puisse continuer de les utiliser pendant longtemps. Citons, par exemple, le Tagged Interchange File Format (TIFF) ("format de fichier d'image en mode point") pour les images, et la famille de formats du Standard Generalized Markup Language (SGML) ("langage de balisage universel normalisé"), qui comprend les langages HTML et le XML pour les documents structurés ;
- là où ces formats largement adoptés se fondent sur des caractéristiques dont le code source n'est pas protégé (comme c'est le cas avec le TIFF et le SGML), il devrait être assez facile de trouver ou de créer des outils lorsqu'ils sont nécessaires pour fournir un accès continu à travers les changements de technologie ;
- une approche concertée vis-à-vis d'un certain nombre de processus de préservation, y compris l'enregistrement des métadonnées, les processus de transfert, la protection des données et l'identification de l'élément. Les normes ne cessent d'évoluer dans ces domaines ;
- la définition des responsabilités et des fonctions de certains types de programmes de préservation. Le grand exemple en est à ce jour le Reference Model for an Open Archival

Information System (OAIS), agréé par l'International Standard Organisation ("l'organisme international de normalisation") en 2002.

On ne saurait considérer les normes comme une panacée pour la préservation. Même là où elles existent, elles sont soumises à des modifications, à une gestion des versions et à une utilisation hors norme par les producteurs. De nombreuses normes de format sont en fait des spécifications protégées qui ne peuvent pas être mises à la disposition du public, aussi peut-il être impossible de savoir si des outils pour un futur accès seront disponibles.

Une normalisation accrue des processus ne pourra qu'aider les programmes de préservation, mais il est important d'entreprendre d'ores et déjà une action de préservation raisonnable sans attendre qu'émerge une "norme de préservation numérique" unique.

### ***10.6.2 Les structures organisationnelles***

Nombre de modèles différents servent comme structures organisationnelles pour gérer la préservation numérique. Parmi ces possibilités figurent :

- la création d'une unité de préservation numérique distincte conçue pour veiller sur tous les aspects ;
- une série d'unités spécialisée veillant sur des aspects distincts ;
- une cellule mère de personnes qui travaillent dans différents domaines et relèvent d'un responsable de programme global ;
- un flux dominant de travail qui traverse des aires de travail déjà existantes si bien qu'il participe au travail courant intégré dans les autres opérations ;
- une insertion du programme dans une aire de travail existante spécifique comme l'IT (la technique de l'information), la préservation, l'élaboration de la collecte, ou les sections de gestion de la collecte.

Ces différents modèles tendent à produire des modes d'accentuations différents en reflétant différents niveaux de confort à l'aide des outils IT (de la technique de l'information), en réunissant des objectifs de collecte, en forgeant une réflexion sur la préservation, etc. Chaque modèle peut devenir efficace, à condition qu'il s'appuie sur les perspectives et les compétences nécessaires et qu'il soit soutenu par une gestion ferme.

### ***10.6.3 La politique et la planification de préservation***

Il convient de mener les programmes de préservation dans le cadre d'une politique qui précise les objectifs du programme et la manière de les atteindre. Dans un domaine d'une telle complexité et dont la compréhensibilité évolue sans cesse, le document qui ose cette politique doit indiquer une direction claire et à longue échéance tout en révisant régulièrement ses orientations.

Lors de la mise en oeuvre de cette politique et de l'élaboration des plans d'action, il faut presque toujours arrêter les questions, les actions et les matériaux dont on se préoccupera en premier lieu, et déterminer le travail à effectuer en priorité.

Certaines questions aident généralement à établir des priorités :

- Qu'est-ce qui est le plus important pour soutenir ou remplir les responsabilités prises par le programme (y compris les exigences juridiques) ?
- Qu'est-ce qui est le plus menacé ?
- Quel est le matériau le plus susceptible d'être demandé mais également de devenir inutilisable ?
- Quelles sont les menaces contre lesquelles il sera le plus facile d'agir ?
- Quelle action faciliterait la vie si l'on s'en occupait immédiatement et qu'est-ce qui rendrait la vie plus difficile si l'on ne s'en occupait pas ?

#### ***10.6.4 Les prestataires de service***

Il peut y avoir des moyens de remplacement pour remplir les engagements de préservation du programme. Les organisations doivent étudier si elles obtiendront un meilleur résultat en faisant le travail en interne, ou en sous-traitant le travail avec quelqu'un d'autre, ou en conjuguant plusieurs méthodes. Plusieurs processus, tels le stockage et la documentation, peuvent faire l'objet d'une sous-traitance : d'où une possibilité d'accès à des compétences et à des équipements spécialisés dont on ne pourrait disposer autrement. Dans bien des cas, les programmes risquent de ne pouvoir apporter les investissements de départ qu'impliquent l'installation et l'entretien de l'infrastructure : payer quelqu'un d'autre pour fournir des services présente alors un intérêt certain.

La sous-traitance s'accompagne de certains risques potentiels, parmi lesquels :

- la création d'une distance entre les objectifs du programme et la prestation de service ;
- l'éventualité d'être prisonnier des services du prestataire parce que rechercher d'autres solutions coûte plus cher que continuer ;
- un contrôle moindre au jour le jour ;
- des coûts plus élevés (ils peuvent aussi être moins élevés) que des dispositifs internes sur le long terme ;
- moins d'occasions d'acquérir un savoir qui permette de s'orienter vers une meilleure pratique.

La réussite de la sous-traitance dépend souvent de la capacité du responsable de programme à définir ce qui est nécessaire ; du choix d'un fournisseur fiable qui puisse offrir les services demandés à un prix convenable ; de l'aptitude à négocier un contrat adéquat et offrant les garanties requises ; et des compétences dans la gestion du contrat.

Dans le travail avec des prestataires de service, les responsables de programme doivent veiller à :

- comprendre leur travail et ce qu'ils doivent accomplir ;
- communiquer leurs besoins ;

- évaluer rigoureusement les aptitudes, les motivations et les connaissances des fournisseurs potentiels ;
- préparer et négocier un contrat adéquat ;
- garantir que les voies de transmission sont ouvertes afin que le moindre problème soit rapidement signalé ;
- surveiller étroitement le déroulement des opérations et jauger régulièrement la bonne marche des accords ;
- s'assurer que des stratégies de secours et des plans de remplacement fiables sont prêts à être mis en oeuvre lorsque les accords prendront fin.

## REFERENCES - pour en savoir plus

### Renvois internes

*La portée des programmes*, voir également Assumer une responsabilité : chapitre 9.

*Garder le matériau accessible*, voir aussi Maintenir l'accessibilité : chapitre 17.

*L'équipement*, voir aussi Protéger les données : chapitre 16.

*Les normes*, voir aussi Travailler avec les producteurs : chapitre 13, et Maintenir l'accessibilité : chapitre 17.

### Renvois externes (tous les documents en ligne ont été consultés à la date de mars 2003)

- Committee on International Cooperation, University Archivists Group (UAG), (2001), *Standards for an Electronic Records Policy*. <http://www-personal.umich.edu/~deromedi/CIC/cic4.htm>
- International Council for Scientific and Technical Information (ICSTI), (2002), *Digital Archiving: Bringing Issues and Stakeholders Together - An Interactive Workshop Sponsored by ICSTI et ICSU Press, 30-31 January 2002*. Maison de l'UNESCO, Paris. <http://www.icsti.org/2000workshop/index.html>
- Jones M., Beagrie N., (2001), *Preservation Management of Digital Materials: A Handbook*, The British Library, Londres.
- Lawrence G. W., Kehoe W. R., Rieger O. Y., Walters W. H., Kenney A. R., (2000), *Risk Management of Digital Information: A File Format Investigation*, Council on Library and Information Resources, Washington, D.C. <http://www.clir.org/pubs/reports/pub93/contents.html>
- Bibliothèque nationale d'Australie, (2002), *A Digital Preservation Policy for the National Library of Australia*. <http://www.nla.gov.au/policy/digpres.html>
- Price L., Smith A., (2000), *Managing Cultural Assets from a Business Perspective*. Council on Library and Information Resources, <http://www.clir.org/pub/reports/pub90/contents.html>

## Chapitre 11. Travailler ensemble

### INTRODUCTION

#### 11.1 Les objectifs

Le propos de ce chapitre est d'inciter les responsables de programme à voir dans la collaboration un moyen de réaliser leurs objectifs de préservation, et de fournir des informations de base sur les divers choix qui peuvent aider à déterminer les modèles à suivre les plus adaptés.

#### 11.2 En bref

Il y a pour les programmes de préservation de bonnes raisons technologiques, économiques et politiques de coopérer. Les décisions relatives à une collaboration doivent se fonder sur une évaluation des bénéfices attendus et des dépenses à engager. Où chercher des partenaires, sur quel point précis portera la relation et quels seront les encadrements structurels appropriés ? Un certain nombre de possibilités sont à envisager. Une coopération réussie résulte habituellement d'un examen approfondi de ces choix et du plein accomplissement de l'effort requis pour gérer pratiquement la collaboration.

### PERSPECTIVE DE GESTION

#### 11.3 La nécessité de collaborer

Patrimoine numérique et collaboration s'accordent fort bien. La technologie des matériaux numériques aide à la collaboration : de tels matériaux sont faciles à dupliquer et nombre d'entre eux sont conçus pour être mis en réseau, aussi la gestion à distance n'est-elle pas difficile.

Il est également coûteux - souvent trop coûteux - d'installer la totalité de l'infrastructure de préservation numérique pour tous les programmes de préservation, de sorte que l'on est fortement incité à trouver des moyens de partager les installations.

On trouve un impératif politique à travailler en coopération : la communauté espère à bon droit que des programmes collaborent pour assurer la préservation de la plus grande part possible de patrimoine numérique, et de façon aussi cohérente que possible.

#### 11.4 Les avantages potentiels de la collaboration

La collaboration a un coût. Négocier des accords, travailler avec des partenaires à distance et maintenir la dynamique prend du temps et de l'énergie. Les priorités organisationnelles peuvent se trouver mises de côté à la suite de problèmes qui surviennent dans la relation de collaboration elle-même, et détournent l'attention de la véritable mission, qui est de préserver les matériaux numériques. Devant ces coûts et ces difficultés potentiels, il est urgent d'identifier les avantages que toute collaboration est censée procurer.

Les avantages qui incitent à entreprendre un effort de coopération comprennent notamment :

- l'accès à un éventail plus large de compétences
- des coûts de développement partagés ;

- l'accès à des outils et à des systèmes qui risqueraient autrement d'être indisponibles ;
- des occasions de formation partagées ;
- une couverture accrue des matériaux préservés ;
- une meilleure planification pour réduire le gaspillage des efforts ;
- un encouragement pour d'autres parties prenantes influentes à prendre la préservation au sérieux ;
- une influence partagée sur les accords avec les producteurs ;
- une influence partagée sur la recherche et le développement des normes et des pratiques ;
- le fait d'attirer des ressources et autres soutiens pour des programmes bien coordonnés sur les plans régional, national ou sectoriel.

### **11.5 Comment parvenir aux avantages d'une collaboration ?**

D'ordinaire, les avantages d'une collaboration n'arrivent pas par hasard, mais résultent d'une étude attentive des choix. Il faut que les programmes prennent en compte leurs partenaires potentiels, les manières de travailler ensemble, les modèles structuraux et les mesures impliquées par la réalisation du travail de collaboration choisi. (Enfin, ils peuvent aussi devoir prendre en considération les occasions proposées.)

#### ***11.5.1 Les partenaires***

Les partenaires potentiels peuvent être des personnes qui travaillent dans le même secteur, qu'il s'agisse d'un consortium de bibliothèques universitaires, ou d'un réseau d'archives de données, ou d'organismes gouvernementaux qui acceptent d'utiliser le même logiciel d'application, ou encore d'un groupe de studios d'enregistrement qui partagent les dispositifs de stockage pour les originaux. On peut se joindre à une collaboration existante, ou constituer de nouveaux partenariats.

Toutefois, il peut être aussi avantageux de regarder au-delà des frontières sectorielles, d'autant que les technologies numériques et les attentes des utilisateurs estompent de plus en plus les limites entre les secteurs. Par exemple, un certain nombre de bibliothèques, d'archives, d'instituts de recherche, d'archives de données et de producteurs d'une zone régionale envisageront d'unir leurs forces pour élaborer un programme local qui réponde à tous leurs besoins.

Il se présentera aussi des occasions d'accords de coopération formels entre des programmes de préservation et les parties prenantes intéressées, notamment des groupes de producteurs, des instances professionnelles, des groupes d'utilisateurs, des groupes de l'industrie de la technique de l'information, ou des organismes gouvernementaux qu'intéresse le développement d'une bonne pratique.

#### ***11.5.2 Les manières de travailler ensemble***

Les programmes de préservation peuvent travailler ensemble de bien des manières, selon les avantages qu'ils veulent obtenir et ce que chaque partenaire a à offrir. Elles comprennent :

- *des normes partagées* : des accords pour faire les choses de la même manière, soit en vue d'un interfonctionnement entre les programmes, soit en se fondant simplement sur une appréhension commune des pratiques qui seront les plus aptes à soutenir les objectifs de préservation.
- *le partage de l'information* : des accords pour partager l'information soit à un niveau général, soit sur des points spécifiques comme des caractéristiques d'approvisionnement, soit des résultats de recherches.

- *parler d'une même voix* : des accords pour élaborer et proposer un message commun à transmettre dans les campagnes de soutien, ou dans une publicité visant à rehausser l'image de la préservation du patrimoine numérique.
- *la division du travail* : des accords pour travailler ensemble à un échelon opérationnel, en prenant des mesures de préservation d'une manière coordonnée, les responsabilités étant assumées soit en parallèle soit réparties entre les programmes.
- *des ressources partagées* : des accords pour partager les ressources comme les systèmes, le personnel ou les capitaux afin de travailler sur un programme commun.

### **11.5.3 Les modèles structuraux**

La plupart des collaborations semblent correspondre à l'une des quatre catégories de modèles structuraux, chacun d'eux présentant des points forts et des faiblesses différentes :

- *des modèles centralisés répartis* : un seul partenaire conduit la politique, fixe les orientations et fournit la majeure partie de l'infrastructure ; il travaille avec un certain nombre d'autres personnes qui ont des rôles clairement définis mais limités, comme l'identification du matériau à préserver et le rajout de métadonnées, éventuellement avec une responsabilité limitée pour un entretien sur le long terme. (Par exemple, un service central d'archives qui travaille avec des organismes d'affaires publics, qui fixe des normes et fournit une orientation.)

Comme tous les modèles répartis, celui-ci offre un partage des coûts et crée un ensemble d'idées et de perspectives. Il permet des économies d'échelle si des fonctions comme le stockage sont centralisées. Il offre une préservation plus sûre parce que l'on contrôle mieux les processus et que l'on utilise des compétences plus spécialisées que dans d'autres modèles. La prise de décision, en grande partie du ressort de l'organisation centrale, se révèle plus efficace que dans des modèles plus également répartis.

En revanche, ce modèle risque de ne pas encourager la propriété du programme chez les partenaires périphériques, car il ne favorise pas le transfert des compétences depuis l'organisation centrale.

Ce type de modèle est probablement efficace pour les programmes débutants, qui essaient de collaborer avec des programmes vastes et perfectionnés. Il convient également, lorsqu'un programme est prêt à assumer une responsabilité durable et qu'un certain nombre d'autres peuvent l'aider mais sans être certains de pouvoir s'engager sur le long terme.

*Des modèles plus également répartis* : un certain nombre de partenaires opèrent à même niveau d'engagement et de responsabilité. (Par exemple : un groupe d'archives de données qui décide de se mettre d'accord sur les normes et de partager les spécifications pour l'achat d'un équipement informatique.)

Ce modèle offre aussi un partage des coûts et la mise en commun d'idées, mais il offre l'avantage d'encourager des niveaux partagés de propriété, sans qu'un des partenaire ait à supporter la pression de prendre seul les décisions. D'un autre côté, il peut être difficile de mettre en place une direction efficace, et les consultations et la prise de décisions risquent de gaspiller du temps. Les économies d'échelles risquent également d'être perdues si des systèmes centralisés importants sont remplacés par un certain nombre de petits systèmes parallèles.

Ce modèle convient sans doute lorsqu'il existe un certain nombre d'acteurs prêts à partager les responsabilités, mais sans qu'aucun ne veuille diriger un programme.

*Des collaborations très largement réparties* : un grand nombre de partenaires jouent chacun un rôle très limité, parfois cantonné à son archivage personnel. (Par exemple : les réseaux des projets d'une communauté locale qui décide de conserver tous ses matériaux pour la postérité.)

Ce modèle sera peut être un point de départ utile pour un programme de préservation : il sensibilisera et permettra de prendre quelques mesures. Cependant, il ne peut pas offrir une grande fiabilité sans un investissement important dans les spécifications, la formation et le contrôle. Cela risque d'entraîner de fortes dépenses globales, même si le modèle, en raison des faibles coûts qu'il représente pour lui, séduit chacun des partenaires. Il sera peut-être malaisé pour ces modèles de traiter les problèmes de préservation sur le long terme d'une manière coordonnée.

Un tel modèle est à recommander lorsqu'il existe un certain nombre de petits sites capables de prendre une responsabilité limitée, notamment si l'un des partenaires peut jouer un rôle de coordination. Il fonctionne également pour un matériau dont la préservation est souhaitable plutôt qu'essentielle.

*Les dispositifs non mis en réseau* favoriseront une collaboration ultérieure en permettant aux programmes de développer des compétences, des stratégies et des systèmes avant de chercher des partenaires appropriés. Les programmes qui fonctionnent dans un environnement où il n'y a pas de partenaires potentiels appropriés peuvent cependant connaître un bon développement de leur côté et trouver, le cas échéant, des occasions de collaboration.

(Par exemple : un service de recherche fonctionnant dans une nouvelle discipline et dans un lieu isolé décidera qu'il doit préserver ses données et monter un programme modeste pour documenter, sauvegarder et transférer celles-ci, en espérant trouver par la suite un programme national ou international qui le prendra sous sa responsabilité.)

#### ***11.5.4 Instaurer une collaboration***

L'expérience montre que les organisations travaillent souvent ensemble avec succès lorsqu'elles :

- comprennent clairement ce qu'elles veulent réaliser en collaboration ;
- choisissent des partenaires appropriés qui sont en mesure de collaborer
- partagent des intérêts et un engagement définis par des discussions et confirmés par l'action
- affectent des ressources suffisantes pour remplir leurs engagements : il est difficile de poursuivre une collaboration dans une ambiance de frustration et d'échec ;
- communiquent souvent et avec efficacité, à la fois à l'échelon opérationnel voulu et par l'intermédiaire d'une sorte de comité de gestion pour le programme collectif ;
- fixent des objectifs réalistes et évaluent régulièrement les dispositifs.

## REFERENCES - **pour en savoir plus**

### **Renvois internes**

*Les coûts*, voir aussi Les coûts de ressources : chapitre 10.

*Les normes*, voir aussi Des outils utiles pour les responsables : chapitre 10.

### **Renvois externes - études de cas** (tous les liens ont été consultés à la date de mars 2003)

On dispose de peu d'analyses de collaboration dans le domaine de la préservation numérique (en revanche il en existe de nombreuses qui abordent la collaboration dans d'autres aspects de la gestion de l'information numérique). La liste suivante est donc centrée sur un éventail de projets et de programmes qui méritent d'être étudiés.

De nombreuses collaborations de préservation numériques réputées sont à l'origine des projets de recherche et des études de faisabilité. Celles qui sont bien documentées comprennent :

- CEDARS, une collaboration entre trois bibliothèques universitaires du Royaume-Uni, a élaboré un système prototype d'archivage numérique décentralisé, d'une pertinence particulière, en même temps que des rapports importants. <http://www.leeds.ac.uk/cedars/>
- CAMiLEON, un projet de recherche mené conjointement par l'Université du Michigan (Etats-Unis) et l'Université de Leeds (Royaume-Uni), qui étudie les méthodes permettant de maintenir l'accessibilité ; <http://www.si.umich.edu/CAMILEON/>
- NEDLIB, une collaboration européenne entre neuf bibliothèques nationales, une institution d'archives nationales et trois grands éditeurs, qui a produit un certain nombre d'outils adaptés aux programmes décentralisés. <http://www.kb.nl/coop/nedlib/>

Parmi les nombreux programmes de préservation actifs construits sur des modèles de collaboration diverses figurent notamment :

- l'Austrian On-Line Archive (AOLA), une initiative conjointe de la Bibliothèque nationale d'Autriche et du Département de technologie du logiciel de l'Université technique de Vienne. AOLA est un service d'archivage d'instantanés de l'espace autrichien de la Toile. <http://www.ifs.tuwien.ac.at/~aola/>
- l'Academic Research in the Netherlands Online (ARNO) qui relie les serveurs documentaires de l'Université d'Amsterdam, de l'Université Tilburg et de l'Université de Twente pour rendre et maintenir accessibles électroniquement leurs travaux universitaires. <http://www.uba.uva.nl/en/projects/arno/>
- l'Australian Digital Theses Project qui vise à établir une base de données distribuée et conservée des versions numériques des thèses réalisées par les étudiants de la recherche de 3e cycle dans les institutions participantes. <http://adt.caul.edu.au/>
- le Projet de Bibliothèque numérique de Chine qui envisage de créer un centre d'archivage de données numériques coordonné par la Bibliothèque nationale de Chine ;
- la "Purge Alert", une initiative internationale du Committee on Earth Observation Satellites (CEOS), destinée à encourager les membres de la communauté des données spatiales globales à transférer la responsabilité d'ensembles de données encore précieuses

avant qu'elles ne soient effacées par leurs conservateurs d'origine. [http://edc.usgs.gov/archive/ceos/data\\_purge\\_alert.html](http://edc.usgs.gov/archive/ceos/data_purge_alert.html)

- le Digital Image Archive of Medieval Music (DIAMM), un projet commun à l'Université d'Oxford, au Royal Holloway et à l'Université de Londres, en accord avec le Service de données des arts et des humanités, qui constitue un système permanent d'archives électroniques de la musique polyphonique médiévale européenne. <http://diamm.ac.uk/>
- le Digital Imaging Project of South Africa (DISA), une collaboration nationale des principales institutions de recherche d'Afrique du Sud, qui fonctionne comme un gisement numérique fiable sur le modèle de l'OAIS et se déroule dans un cadre d'accords formels entre les participants. <http://disa.nu.ac.za/nu.ac.za>
- l'European Visual Archive (EVA), qui est centrée sur la fourniture d'un accès facile et préservé aux informations et aux collections intégrées que conservent les archives européennes. <http://www.eva-eu.org/>
- JERRI: Joint Electronic Records Repository Initiative de l'Ohio. Ce projet mené conjointement par l'Ohio Historical Society, la State Library of Ohio, l'Ohio Supercomputer Center et l'Ohio Department of Administrative Services, est destiné à maintenir l'accès public aux archives électroniques d'Etat et aux publications originaires du Web de valeur historique durable via des archives électroniques. Il fonctionne en partenariat avec le projet de gestion et de préservation des collectes numériques de l'OCLC. <http://www.ohiojunction.net/jerri/>
- le National Digital Archives programme (NDAP), lancé à Taiwan en 2002, fait collaborer neuf organisations nationales qui comprennent des musées, des bibliothèques, des archives, des institutions universitaires et le gouvernement.
- les archives radiophoniques numériques norvégiennes, une collaboration entre la Bibliothèque nationale de Norvège et la Radiodiffusion norvégienne qui vise la création d'archives communes pour gérer et préserver des quantités importantes de fichiers audio-numériques.
- PANDORA, un programme lancé par la Bibliothèque nationale d'Australie en partenariat avec les bibliothèques de l'Etat et du territoire australiens et les archives de l'écran et du son, ScreenSound (Australie), pour saisir, préserver et fournir l'accès aux publications australiennes en ligne. <http://pandora.nla.gov.au/>
- la Victorian Electronic Records Strategy (VERS), un projet entrepris par le Public Record Office Victoria (PROV) de concert avec l'Australian Commonwealth Scientific and Industrial Research Organisation (CSIRO) et Ernst et Young en 1998. Le projet a réalisé une stratégie destinée à être utilisée par les organisations gouvernementales de Victoria pour la préservation à long terme des enregistrements électroniques. <http://www.prov.vic.gov.au/vers/>

**TROISIEME PARTIE**

**PERSPECTIVES TECHNIQUES ET PRATIQUES**

## Chapitre 12. Décider de ce que l'on conserve

### INTRODUCTION

#### 12.1 Les objectifs

Ce chapitre permettra au lecteur de mesurer l'enjeu majeur que représente le choix même des matériaux numériques à sélectionner pour la préservation et de bien comprendre quelques principes d'orientation. Il donne aussi des conseils pratiques et techniques, notamment des propositions sur la manière d'identifier les éléments essentiels à préserver.

#### 12.2 En bref

Il est habituellement nécessaire de décider quels matériaux numériques méritent d'être conservés, comme il en va pour les matériaux non numériques. De nombreuses méthodes similaires - une sélection fondée sur des critères contenus dans les politiques de développement de la collecte, ainsi qu'une bonne connaissance des matériaux et de leur contexte - sont fondamentales dans la sélection du patrimoine numérique. Les programmes de préservation ont également besoin de définir les éléments ou caractéristiques qui donnent une valeur aux matériaux qu'ils sélectionnent, de sorte que ces éléments puissent être conservés.

#### 12.3 La terminologie

Le mot *sélection* a été employé dans ce chapitre en tant que terme générique. Il faut comprendre qu'il englobe des concepts comme l'*appréciation* qui a un sens particulier pour la communauté des services d'archives.

### LES PRINCIPAUX PROBLEMES DE GESTION

#### 12.4 Ce que l'on doit garder

Décider ce qui doit être préservé, par qui, et pour combien de temps, telles sont les décisions fondamentales à prendre dans la gestion des différentes sortes de patrimoines matériels. Ces décisions sont indispensables car les choses à garder sont généralement bien plus nombreuses - plus d'informations, plus d'enregistrements, plus de publications, plus de données - que les moyens dont on dispose à cet effet. Chaque choix de préservation se fait au détriment de quelque chose d'autre.

#### 12.5 Construire sur des concepts existants

La sélection du patrimoine numérique est conceptuellement la même que la sélection des matériaux non numériques. Tout programme existant doté de procédures bien établies d'évaluation et de sélection du matériau pour la préservation aura déjà une politique, des compétences et des outils qui peuvent servir à sélectionner les matériaux numériques, même s'ils nécessitent une adaptation.

#### 12.6 Le défi lancé par les matériaux numériques

Les matériaux numériques lancent des défis d'un type nouveau que les responsables de programme doivent relever en prenant les meilleures décisions possibles en matière de sélection.

- il y a souvent un volume important de matériaux à évaluer ;
- les moyens de production et de diffusion des matériaux numériques sont largement disponibles, aussi leur qualité est-elle souvent inégale ;
- en même temps, une pression peut s'exercer pour préserver la totalité de la circulation qu'empruntent les nouvelles voies de communication comme la Toile mondiale, indépendamment de la qualité ;
- le calendrier de la sélection est en général un problème crucial car les matériaux numériques se désélectionnent rapidement en devenant inutilisables. Il faut parfois décider de faire la sélection sans pouvoir attendre que la valeur durable du matériau soit avérée ;
- certains objets numériques peuvent se révéler difficiles à épingle. Des genres nouveaux risquent de ne pas entrer dans les classifications existantes ; certaines ressources numériques se composent de parties qui sont reliées ou qui se chevauchent ; beaucoup, également, existent en versions parallèles. La sélection doit s'opérer à travers ces complications à clarifier pour parvenir à des décisions sans ambiguïté sur ce qu'il faut préserver ;
- même lorsqu'on a délimité des frontières extérieures, il peut être difficile de savoir quels sont les éléments à conserver pour que l'objet numérique atteigne son objectif principal ;
- il peut également être malaisé de savoir la provenance des matériaux numériques, ce qui rend difficile de déterminer qui est responsable de leur préservation et avec qui négocier les droits qu'exigent les programmes de préservation.

## DES PRINCIPES POUR RELEVER CES DEFIS

### 12.7 Des décisions fondées, cohérentes et transparentes

Les processus de sélection doivent souvent composer avec des incertitudes et ils impliquent des jugements qui sont subjectifs et spéculatifs ; ils n'en doivent pas moins être fondés, cohérents et transparents :

- les décisions doivent être fondées sur une bonne connaissance du matériau, de son contexte et des besoins des parties prenantes ;
- les décisions doivent avoir pour base cohérente une politique de sélection reflétant les objectifs de l'Organisation qui assume la responsabilité de la préservation. Pour les institutions de collecte comme les bibliothèques, les musées et les archives, une politique de développement de collecte existante peut fournir une bonne direction ;
- pour qu'ils soient transparents et que l'on puisse en rendre compte, les processus de sélection doivent être visibles, fondés sur des documents stratégiques accessibles au public, et fournir des arguments clairs et explicites à l'appui de ce qui a été sélectionné et de ce qui a été exclu.

## **12.8 Une base pour des critères de sélection**

Il n'est pas possible de proposer des critères spécifiques pour la sélection de matériaux du patrimoine numérique parce que l'on juge qu'ils méritent d'être conservés pour diverses raisons. Toutefois, en principe :

- les décisions doivent se fonder en premier lieu sur la valeur du matériau dans le soutien qu'il apporte à la mission de l'organisme qui prend des responsabilités de préservation ;
- cette valeur doit être évaluée par rapport aux coûts probables et aux difficultés de la préservation, ainsi qu'à la disponibilité attendue des ressources. Bien des arguments incitent à commencer par des matériaux qui peuvent être facilement sauvegardés. Toutefois, les capacités et les coûts futurs des programmes de préservation numérique étant encore indéterminés, il serait irresponsable de refuser un matériau précieux sous prétexte qu'il semble difficile à préserver ;
- si des programmes de préservation ne sont pas en mesure de gérer un matériau qu'ils estiment bon à préserver, ils doivent l'indiquer dans leurs politiques de sélection ;
- il est souhaitable que l'effet global de tous les efforts de collecte et de préservation préserve au moins un échantillon de chaque sorte de matériaux numériques, y compris des échantillons de ce qui est simplement éphémère.

## **12.9 Reconnaître les éléments qui donnent sa valeur au matériau**

Décider de choisir un élément, ou une classe d'éléments, à préserver peut être insuffisant :

- la préservation implique la conservation des éléments et des caractéristiques qui donnent sa valeur au matériau. Le processus de sélection doit envisager quels sont ces éléments et ces caractéristiques.
- le processus doit documenter les raisons pour lesquelles le matériau a été choisi de sorte que les responsables de la préservation puissent comprendre ce qu'on leur demande de conserver. (Quelques notes plus détaillées figurent à la fin de ce chapitre.)

## **12.10 Une approche prudente**

La décision de ne pas préserver est généralement une décision définitive pour les matériaux numériques. Une approche prudente consiste à décider quels sont les matériaux qui doivent être précisément préservés et pour combien de temps ; et quelle action de préservation provisoire doit-on entreprendre en attendant que soit prise une décision de sélection plus définitive.

## **PROBLEMES TECHNIQUES ET PRATIQUES**

### **12.11 Suppositions sur la valeur**

Il est dangereux de supposer que les estimations qui sont faites actuellement de la valeur des matériaux sont un guide fiable pour estimer la valeur future. Par exemple, d'anciennes données de télédétection collectées au cours des dernières décennies ont acquis une importance inattendue pour évaluer les modifications de l'environnement. Cette expérience suggère qu'il vaut probablement mieux collecter trop de matériaux que pas assez, si le programme de préservation peut le faire.

### **12.12 La documentation**

Lorsque les matériaux numériques ne se comprennent que par référence à un ensemble de règles, tels qu'un système de conservation d'enregistrement, un système de création de base de données ou de données, ou quelque autre information contextuelle, le processus de sélection doit identifier la documentation qu'il faut également préserver.

### **12.13 Le rôle des producteurs**

Les producteurs de matériaux numériques peuvent avoir un rôle important à jouer dans la sélection de ce qui doit être préservé. Ils sont souvent bien placés pour comprendre pourquoi des objets numériques ont été créés, quel est leur "message" essentiel, et quelles relations existent entre les objets et leur contexte. Si cette information n'est pas saisie chez le producteur, il sera peut-être très difficile de la reconstituer ultérieurement.

### **12.14 Une collecte sélective ou exhaustive**

La question du choix de la collecte peut se poser : vaut-il mieux qu'elle soit exhaustive ou sélective ? (Ce point surgit fréquemment dans des discussions concernant des matériaux devenus disponibles sur la Toile mondiale, par exemple.)

Il existe de solides arguments en faveur des deux méthodes. Les avocats d'une méthode exhaustive soutiennent que toute information peut se révéler avoir une valeur à long terme, et que les coûts d'une sélection détaillée sont plus élevés que les coûts de la collecte et du stockage de toutes les informations. Les avocats d'une méthode plus sélective soutiennent que celle-ci leur permet de créer des collectes de ressources de grande valeur, avec l'assurance d'une qualité technique et la possibilité de négocier les droits d'accès avec les producteurs.

Il pourrait y avoir une place pour les deux méthodes, étant donné qu'elles sont susceptibles de produire des collectes totalement différentes de patrimoine numérique qui sont appréciées à des fins différentes.

### **12.15 Les accords de collecte**

Pour minimiser le risque de négliger des matériaux importants, et éviter une multiplication inutile des efforts, il peut être nécessaire de chercher des accords avec d'autres organismes de collecte et de préservation concernant les responsabilités et les rôles respectifs.

### **12.16 Définir les éléments**

Il arrive que les politiques de sélection doivent décider s'il faut sélectionner la totalité des éléments et des collectes, ou seulement des échantillons. Il est généralement préférable de préserver tous les éléments pour conserver leur intégrité, mais il peut être nécessaire de restreindre la collecte à des échantillons représentatifs comme moyen d'attester au moins l'existence de certaines sortes de matériaux.

La politique de sélection peut aussi avoir à considérer si la réutilisation du matériau constitue un nouvel élément qui doit aussi être préservé.

### **12.17 Les problèmes de droits**

Les problèmes de droits peuvent influencer sur les décisions de sélection. Les programmes de préservation sélectionnent souvent des matériaux qui sont toujours sujets à des droits, mais en

général ils ne sélectionneront pas un matériau lorsque les droits sont si restrictifs que l'on ne peut négocier des arrangements pour assurer l'accès à une étape ultérieure. Si le matériau ne peut jamais être rendu disponible pour une utilisation, ou si l'on ne peut prendre les mesures de préservation nécessaires, il n'y a guère de raison pour le sélectionner en tant que patrimoine numérique.

### **12.18 La sélection périodique**

Les décisions de sélection doivent-elles être définitives ? Revoir les décisions de sélection en fonction de périodes de conservation précises est une vieille pratique dans la communauté des archives. Cette méthode se révèle pertinente pour d'autres types de matériaux numériques ainsi que pour les enregistrements : elle permet de vérifier si la valeur du matériau justifie toujours le coût de sa conservation. Par contre, le processus de sélection lui-même est coûteux et devra être répété le moins souvent possible. Fait encore plus important, il faudra expliciter toutes les décisions de sélection sujettes à révision afin d'éviter tout risque d'en inférer une responsabilité de préservation permanente.

### **12.19 Soutenir le processus de sélection**

La sélection exige l'attribution de ressources : des personnes qui disposent de connaissances, de temps, d'installations et d'équipement pour examiner le matériau. Gérer la sélection exige aussi l'élaboration de critères pour l'appréciation. Lorsque la quantité du matériau est si importante qu'il est impossible d'évaluer individuellement les éléments, il peut être nécessaire de créer des classes de matériau que l'on évaluera sur la base d'échantillons représentatifs.

## **CONSIDERATIONS PARTICULIERES**

### **12.20 Choisir les éléments et les caractéristiques essentiels qu'il faut préserver**

Les programmes de préservation fonctionnent souvent comme les agents d'autres parties prenantes : ils assument la tâche de préservation pour le compte de quelqu'un qui veut qu'un matériau soit conservé pour une raison précise. Ce "quelqu'un d'autre" peut être aussi vaste et à facettes multiples que "la nation" ou "l'ensemble de la communauté", qui peuvent mandater un programme pour collecter et préserver une gamme très large de matériaux ; ou être aussi restreint que les membres d'un organisme ou des chercheurs travaillant dans une discipline particulière qui veulent que leur contribution personnelle soit préservée pour un usage ultérieur.

Les besoins de la "communauté" - quelle que soit la définition qu'on lui donne - pour laquelle le matériau est conservé ? susciteront de nombreuses décisions, depuis le choix du matériau à sélectionner, jusqu'au type et au niveau de la documentation qui sera archivée, au degré d'exigence en matière d'authenticité et aux stratégies employées. Par exemple, certains programmes doivent offrir aux utilisateurs la possibilité d'interroger d'anciennes données pour produire de nouveaux résultats, tandis que d'autres ont pour mission de proposer un matériau uniquement accessible en lecture pour empêcher qu'il soit modifié ou manipulé. Certains programmes sont même chargés de permettre aux utilisateurs de faire fonctionner de vieilles simulations, de jouer à d'anciens jeux électroniques, ou de visualiser de l'art numérique d'une manière qui reproduise l'expérience originale plutôt qu'une expérience accélérée telle que les technologies plus récentes peuvent en fournir.

Définir les éléments ou les caractéristiques essentiels, qualifiés également de *propriétés significatives* par certains programmes, n'est pas difficile sur le plan conceptuel, comme le montrent les exemples ci-dessus. Dans certaines circonstances - lorsque les attentes de l'utilisateur sont clairement définies et limitées, et que les matériaux facilement caractérisés sont tous similaires -

définir et encoder les éléments essentiels doit être une tâche simple. Par exemple, un programme peut décider que les utilisateurs d'une vaste collection de messages de courriel ont seulement besoin de voir les éléments que l'on peut définir comme "informations de contenu", tels que le nom et l'adresse de l'expéditeur, le sujet, la date et l'heure, les destinataires et le message, dans une structure normalisée avec le formatage le plus simple possible. Avec cette méthode, des archives gouvernementales peuvent espérer appliquer ce modèle d'éléments essentiels à un très grand nombre d'enregistrements de courriels.

En revanche, certains matériaux sont bien plus difficiles à définir, et les attentes quant à la manière dont ils seront représentés pour être utilisés, notamment à l'intention d'un groupe évolutif d'utilisateurs potentiels, peuvent être si difficiles à définir à l'avance que cela devient presque impossible.

Les méthodes dans ce domaine se développent au fur et à mesure qu'un nombre croissant de personnes sont confrontées aux problèmes de la description, du stockage et de la planification pour représenter des objets numériques dans des collections qui vont en grandissant sur de longues périodes de temps.

Alors que se développent des méthodes de plus en plus élaborées pour définir et décrire les éléments essentiels, les questions suivantes peuvent apporter une aide dans le processus de sélection. (On verra que cela fait vraiment partie du processus d'appréciation entrepris par les responsables d'archives en vue de comprendre les archives qu'ils retiennent pour la sélection.)

- Pour qui ce matériau doit-il être conservé ? Y a-t-il des attentes particulières concernant ce qu'on pourra faire avec le matériau quand il sera représenté ?
- Pourquoi les matériaux méritent-ils d'être conservés ? Qu'est-ce qui leur donne la valeur qui justifie la peine de les préserver ? Cette valeur est-elle associée avec :
  - le témoignage ;
  - l'information ;
  - des facteurs artistiques ou esthétiques ;
  - une innovation importante ;
  - une association historique ou culturelle ;
  - ce qu'un utilisateur peut faire avec le matériau, ou ce qu'il peut faire avec le matériau ;
  - des caractéristiques culturellement significatives.
- La valeur est-elle liée à l'aspect du matériau ? (Serait-il perdu ou altéré de façon importante s'il avait une autre apparence ?)
- La valeur est-elle liée à la manière dont l'objet fonctionne ? (Serait-il perdu si des fonctions particulières lui étaient enlevées ? Ou si des fonctions particulières avaient une vitesse modifiée ou nécessitaient différentes frappes ?)
- La valeur est-elle liée au contexte du matériau ? (Serait-il perdu si les liens incorporés dans le matériau ne fonctionnaient pas ? Ou si un utilisateur ne pouvait plus voir les signes qui reliaient le matériau à son contexte d'origine ?)
- Est-il possible de faire la distinction entre des éléments à l'intérieur de chacun de ses domaines ? Par exemple, des bandeaux publicitaires doivent-ils être considérés comme une part essentielle de l'apparence du matériau ? Certains éléments de navigation ou fonctions de représentation sont-ils nécessaires mais pas d'autres ?

- S'il est difficile de définir ce qu'il faut conserver, il peut être plus facile de prendre en compte l'effet d'un élément qui ne doit pas être conservé, et de chercher des fonctions et des éléments qui ne sont pas vraiment nécessaires.

## POUR LES PROGRAMMES DE PRESERVATION DISPOSANT DE PEU DE RESSOURCES

### 12.21 La sélectivité

Les programmes de préservation qui disposent de peu de ressources doivent encore décider des matériaux dont ils assumeront la responsabilité. Les coûts de la préservation étant liés à la quantité de matériau à gérer, de tels programmes seront sans doute obligés d'être extrêmement sélectifs, en limitant leurs ambitions à une petite quantité de matériau de grande valeur.

Les coûts de la préservation sont également liés à la gamme des problèmes et des formats qu'il faudra gérer, aussi peut-il être judicieux de limiter sérieusement les types de matériaux sélectionnés à un très petit nombre de formats.

### 12.22 La coopération

Passer des accords avec d'autres programmes peut alléger le fardeau de la sélection des matériaux. Alors qu'il faudrait normalement négocier ceux-ci avec d'autres programmes de préservation, il peut y avoir une éventualité de conclure avec les producteurs des accords qui les conduiraient à décider de ce qui doit être collecté et préservé. Le programme de préservation devra toujours être responsable du contrôle de la qualité ainsi que de la formulation des critères selon lesquels le matériau doit être choisi.

### 12.23 Commencer simplement

Les processus de sélection peuvent évoluer avec le temps, en commençant par des décisions simples de choisir des matériaux faciles à collecter et à préserver ("à portée de la main"), et en envisageant de prendre par la suite des décisions plus élaborées s'appliquant à une gamme plus large de matériaux plus difficiles.

## REFERENCES - pour en savoir plus

### Renvois internes

*Les éléments essentiels*, voir aussi Comprendre la préservation numérique : chapitre 7, et Maintenir l'accessibilité : chapitre 17.

*Les problèmes de droits*, voir aussi Gérer les droits : chapitre 15.

**Renvois externes** (tous les liens ont été consultés à la date de mars 2003).

- Projet Cedars (1999). Why Can't We Preserve Everything? Selection Issues for the Preservation of Digital Materials. <http://www.leeds.ac.uk/cedars/documents/ABS01.htm>
- Projet InterPARES (2002). The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project. <http://www.interpares.org/book/index.htm>

- Marcum Deanna B. (2001), "Scholars as Partners in Digital Preservation", in CLIR Issues, n° 20, Council on Library and Information Resources. <http://www.clir.org/pubs/issues/issues20.html-scholars>
- National Library of Australia (mise à jour 2002). Guidelines for the Selection of Online Australian Publications Intended for Preservation by the National Library of Australia. <http://pandora.nla.gov.au/selectionguidelines.html>
- Royal Statistical Society and UK Data Archive (2002). Preserving and Sharing Statistical Information. <http://www.data-archive.ac.uk/home/PreservingSharing.pdf>

## Chapitre 13. Travailler avec les producteurs

### INTRODUCTION

#### 13.1 Les objectifs

Ce chapitre a pour objet d'inciter les responsables de programme à envisager les modalités de travail avec les producteurs du patrimoine numérique et de les aider à s'orienter dans le domaine des normes et des pratiques qui facilitent la tâche de préservation.

#### 13.2 En bref

Le patrimoine numérique est souvent créé sans que l'utilisation et l'accessibilité durables soient prises en compte. Or, il existe des normes et des pratiques précises auxquelles peuvent recourir les producteurs aussi bien pour aider que pour entraver la préservation. Les responsables de programme doivent chercher les moyens d'exercer une influence positive le plus tôt possible sur le cycle de vie du patrimoine numérique. Cela exige souvent d'accepter de travailler avec les producteurs.

#### 13.3 La terminologie

Le mot *producteurs* a été employé dans ce chapitre pour désigner tous ceux qui sont impliqués dans la conception, la création et la diffusion des matériaux numériques avant que ceux-ci n'entrent dans un programme de préservation. Les programmes de numérisation entrent précisément dans la catégorie des "producteurs" dont l'apport numérique doit être géré en vue d'une accessibilité durable par les programmes de préservation.

### LES PRINCIPAUX PROBLEMES DE GESTION

#### 13.4 La "préhistoire" du patrimoine numérique

Les matériaux numériques sont créés par des "producteurs" qui ne s'intéressent pas forcément à une disponibilité sur le long terme : il se peut que la création d'un patrimoine numérique n'entre pas dans leurs intentions. Quant à ceux qui souhaitent fabriquer un matériau de valeur durable, ils risquent de ne pas avoir les connaissances ou les moyens pour le faire, ou d'être gênés par d'autres entraves dans leur environnement de travail.

En l'absence d'une intervention précise, il est peu probable que les matériaux du patrimoine numérique seront automatiquement constitués de façon à diminuer les coûts et à faire tomber les barrières qui s'opposent à la préservation. En réalité, de nombreuses pratiques rendent la préservation beaucoup plus difficile.

### **13.5 Les difficultés d'entente avec les producteurs**

En cherchant à travailler avec les producteurs pour surmonter les obstacles de la préservation, il est probable que les programmes devront relever un certain nombre de défis :

- dans bien des cas, le "producteur" est un concept comportant plusieurs couches : il est constitué d'un certain nombre d'agents qui accomplissent des fonctions très différentes comme les développeurs de logiciels, les créateurs (souvent multiples), les rédacteurs, les éditeurs et les prestataires de service ;
- certains producteurs peuvent être réticents, voire hostiles, à l'idée qu'un tiers s'intéresse d'une manière ou d'une autre à la "gestion en tant que patrimoine numérique" de matériaux qu'ils ont créés.

#### DES PRINCIPES POUR RELEVER CES DEFIS

### **13.6 La nécessité de travailler avec les producteurs**

Les efforts de préservation qui attendent jusqu'à ce que des problèmes commencent à apparaître risquent d'être plus onéreux, plus ardues et moins efficaces que les efforts entrepris tôt.

L'expérience a appris aux organismes qui ont à la fois des fonctions de création et de préservation de patrimoine que les soins apportés dès le départ dans l'utilisation des normes, de la documentation, de la bonne gestion des fichiers et autres pratiques, portent leurs fruits ultérieurement en abaissant les coûts de préservation et d'entretien et en facilitant l'accessibilité, la réutilisation et la gestion des collections.

Si tous les programmes de préservation n'ont pas les mêmes possibilités pour influencer les pratiques de production, ils doivent tous chercher à influencer la manière dont les matériaux sont créés, puis gérés, le plus tôt possible dans leur cycle de vie.

### **13.7 Ce que signifie "travailler avec les producteurs"**

En général, travailler avec les producteurs inclut tout ou partie des opérations suivantes :

- les mettre au courant de l'existence, de la mission et des activités du programme ;
- discuter des circonstances dans lesquelles le processus de production peut aider ou entraver le processus de préservation ;
- identifier les avantages respectifs pour les deux parties en réduisant le plus possible les obstacles à la préservation ;
- chercher des moyens réciproquement acceptables de faciliter le processus de préservation ;
- identifier les préoccupations des producteurs et chercher des moyens réciproquement acceptables d'y répondre ;
- s'il convient, donner des conseils détaillés sur les bonnes pratiques telles que l'utilisation de normes, de formats, de gestion de fichier et de métadonnées ;
- négocier des accords pour les transmissions et la gestion des droits ;

- établir des accords pour entreprendre une action spécifique, souvent fondée sur un travail qui passe par des projets pilotes et des évaluations conjointes.

### **13.8 Une collaboration efficace**

L'efficacité d'une collaboration entre un programme de préservation et des producteurs peut dépendre d'une série de facteurs comme :

- la nature de leur relation. Considérer par exemple la différence de poids potentiel pour :
  - un service d'archives organisationnel d'enregistrements doté d'une compétence juridique en matière de création d'enregistrements au sein de l'organisme ;
  - des archives de données reconnues sur le plan national qui négocient avec des chercheurs indépendants qui produisent des ensembles de données dans une discipline universitaire étendue ;
  - des archives audiovisuelles gouvernementales qui cherchent à convaincre des producteurs d'enregistrements indépendants que leurs enregistrements "maison" font partie du patrimoine national ;
  - une petite collection spécialisée qui essaie de préserver des jeux informatiques produits dans le commerce et mis sur le marché international ;
- l'empressement des producteurs à participer ;
- la compétence et les aperçus techniques que le programme de préservation peut proposer ;
- l'habileté du programme de préservation à négocier des accords mutuellement avantageux.

Le programme de préservation doit chercher à optimiser l'efficacité de son influence dans des limites réalistes.

### **13.9 Les avantages**

Travailler avec les producteurs pour surmonter les obstacles de la préservation implique de nombreux avantages potentiels pour le programme de préservation ainsi que pour le producteur. Certains d'entre eux sont présentés dans le tableau 13.1

#### **Avantages à court terme pour le programme de préservation**

L'établissement de points de contact peut faciliter la communication.

Les transmissions peuvent être facilitées, notamment là où la "collecte" automatisée ne fonctionne pas (voir chapitre 14).

L'implication du producteur dans la décision de ce qui doit être préservé.

La compréhension du type de matériau disponible et de la manière dont il est considéré par la communauté du producteur.

L'identification de matériaux autrement "invisibles".

Des transmissions moins coûteuses.

### **Avantages à long terme pour le programme de préservation**

Une amélioration dans le choix des formats et dans la manière de les utiliser, et des possibilités de négocier des accords pour les dispositifs de sécurité de dérivation qui entravent la copie de préservation.

Une transmission améliorée de la documentation.

Une meilleure compréhension des rôles et des responsabilités.

Un aperçu des tendances futures dans la production des matériaux du patrimoine numérique.

Une base pour identifier les problèmes prioritaires avec les communautés spécifiques.

Une préservation moins coûteuse sur le long terme.

### **Avantages pour le producteur**

Une représentation améliorée de sa production dans les collectes archivées.

Un déroulement des opérations plus efficace, moins de travail "refait" pour répondre aux exigences de l'archivage.

Une plus grande reconnaissance de la valeur de son travail.

Le fait que son travail soit porté à la connaissance d'un public plus large peut créer de nouveaux marchés et susciter un plus grand intérêt.

Peut aider à établir une crédibilité pour de nouvelles formes de production et de diffusion de l'information.

Un intérêt accru pour l'utilisation d'un logiciel au code source libre (à des fins de préservation) peut encourager de nouveaux modèles de production en collaboration.

*Tableau 13.1 - Quelques avantages potentiels de la collaboration entre les programmes de préservation et les producteurs de matériaux du patrimoine numérique.*

## **PROBLEMES TECHNIQUES ET PRATIQUES**

### **13.10 Reconnaître les différences**

Il est important de reconnaître que les créateurs de matériaux numériques travaillent dans des environnements différents et sont susceptibles d'être très différents et de bien des manières : la façon dont ils abordent leur travail, la dimension de leurs opérations, l'aide organisationnelle et technique dont ils disposent, et l'intérêt qu'ils portent aux problèmes d'accès à long terme. Par exemple, des scientifiques qui collectent des données s'attacheront particulièrement à une protection précise et sûre de leurs données ; à la façon dont les formats et les normes s'adaptent aux nécessités de leur travail ; à la commodité des accords de transmission ; et au maintien des droits moraux et des contrôles d'accès sur les données. En revanche, les éditeurs commerciaux

d'emballages de CD-ROM s'intéresseront davantage aux contrôles des copies non autorisées ; aux coûts et aux risques liés à la fourniture de versions "non protégées" à un programme de préservation ; à la réutilisation potentielle de leur contenu ; et à leurs obligations de licence envers les propriétaires de logiciels dont ils ont utilisé les produits

### **13.11 Des méthodes pour travailler avec les producteurs**

Il y a pour les responsables de la préservation du patrimoine numérique de nombreuses manières d'aborder une collaboration avec ceux qui le créent et le diffusent.

La première étape est évidemment d'identifier qui est concerné. Telle action pourra utilement être entreprise avec les groupes professionnels représentatifs, tandis que telle autre réclamera un contact et une négociation individuels.

Les créateurs ont également besoin de savoir avec qui ils traitent. Les programmes de préservation doivent s'employer activement à sensibiliser à leur rôle propre.

Il peut être avantageux d'identifier des groupes particuliers de producteurs et de travailler avec eux, en abordant des problèmes spécifiques, plutôt qu'en essayant de résoudre les problèmes de chacun à un niveau général.

A un niveau professionnel plus large, il peut être utile d'élaborer un code d'usage qui fixe les accords conclus sur les rôles et les responsabilités, et définit le champ et les termes de la coopération en cours.

De nombreux secteurs possèdent des groupes professionnels actifs qui constituent autant de forums de discussion possibles. De même qu'ils offrent de occasions de dialogue avec les responsables professionnels, ces forums peuvent aider à établir de nouveaux schémas de réflexion qui intègrent une perspective à plus long terme.

Il est important pour les programmes de préservation de proposer un encouragement et une rétroaction positive pour les mesures que souhaitent prendre les producteurs, et de garantir un certain niveau de transparence en rendant des comptes sur la manière dont les programmes de préservation traitent leurs matériaux. En prouvant que la coopération entraîne une action de préservation efficace, on a une chance d'encourager les producteurs à accepter et à soutenir une collaboration ultérieure.

### **13.12 Une voie à "double sens"**

Dans bien des cas, travailler avec les producteurs implique une véritable contribution de la part du programme de préservation, et pas seulement de celle du producteur. Les domaines de cette contribution comprennent notamment :

- la fourniture de directives écrites et de spécifications ;
- l'apport d'une formation pour le personnel ;
- une aide dans la conception des systèmes et du déroulement des opérations ;
- un échange d'informations et des astuces de travail ;
- des dispositifs de succession pour un matériau figurant dans un programme de préservation géré par le producteur.

## CONSIDERATIONS PARTICULIERES

### 13.13 Spécifications et conseils pour une meilleure pratique

Ces conseils pour une bonne pratique comportent, entre autres :

- des points d'organisation qui faciliteront la gestion des matériaux numériques ;
- une planification du projet, qui met l'accent sur la conception du système avant toute création d'enregistrement ou de publication ;
- le choix d'un support. Il convient d'encourager les producteurs à utiliser des produits "professionnels" qui survivront suffisamment longtemps pour que les données puissent être transférées sur d'autres supports, soit par le producteur, soit par transfert sur le programme de préservation ;
- l'utilisation de formats de fichier et de normes de données adéquats. Sauf s'il y a de très bonnes raisons de faire autrement, il faut encourager les créateurs à utiliser des formats de fichier largement adoptés et très normalisés qui correspondent à leurs objectifs. En général, il est plus facile de préserver des données dans des formats plus simples qui utilisent des logiciels non déposés au code source libre (même si certaines applications brevetées en font une utilisation si étendue qu'on pourrait les considérer comme une norme dans leur domaine, notamment lorsque leurs spécifications sont librement publiées). Les matériaux en ligne publiés pour un accès au public doivent être lisibles par des logiciels de navigation d'usage courant. Pour un matériau de valeur durable, il faut envisager la structuration des documents dans un format normalisé, facilement reconnu et à longue durée de vie comme le XML (Extensible Markup Language) ;
- la validation des formats. Il ne suffit pas de simplement choisir un format normalisé, puis de l'utiliser de façon non normalisée : il faut que les formats soient mis en oeuvre conformément à leur norme et si nécessaire validés pour supprimer toute particularité susceptible de compliquer la préservation. (Il existe de nombreux outils en ligne pour valider toute une gamme de formats de fichier ;)
- les noms de fichier doivent être cohérents et sans ambiguïté ;
- les fichiers en ligne doivent être gérés pour un accès continu par l'intermédiaire d'un service d'identification et de conversion, ou de réacheminement de messages si les fichiers sont déplacés. Un certain nombre de systèmes IP sont utilisés internationalement dans divers secteurs, même si aucun n'est utilisé universellement. Le système DOI (Digital Object Identifier) utilisé par les éditeurs commerciaux pour gérer les droits est le plus largement accepté ;
- les créateurs doivent créer des métadonnées de bonne qualité pour les ressources qu'ils créent, en utilisant un système largement accepté comme le MARC ou les métadonnées du Dublin Core, par exemple. Les métadonnées aideront les utilisateurs à trouver et à utiliser leurs ressources. Il faut également enregistrer les métadonnées qui décrivent la nature technique des objets numériques, ce qui est exigé pour y accéder et tous les changements dans ces détails au cours de leur cycle de vie : ces informations seront nécessaires dans leur gestion. Les métadonnées peuvent être soit incorporées dans les ressources, soit stockées dans un fichier de métadonnées relié ;

- la gestion de fichier. Les fichiers originaux de préservation doivent être stockés et gérés séparément des copies de diffusion. Les procédures de gestion des bases de données doivent garantir que les données ne soient pas écrasées avant d'être saisies ;
- la sécurité du système. Les fichiers et les systèmes doivent être totalement protégés des détériorations ou de la perte par l'adoption de meilleures mesures de sécurité pratiques et par des dispositifs de sauvegarde adéquats même pour un stockage à court terme ;
- l'authenticité. Tous les fichiers doivent être identifiés et leur origine comme leur histoire documentées pour fournir une preuve durable de leur authenticité ;
- la formation. Le personnel, les fournisseurs et tous ceux qui entrent en relation avec les matériaux numériques doivent être guidés par des procédures et des manuels adéquats ; il faut également qu'ils soient formés, motivés et équipés de manière adéquate pour les utiliser ;
- si des obstacles à l'accès ou à la copie sont jugés nécessaires pour protéger la propriété intellectuelle, ils peuvent rendre la préservation impossible. Il faudra des accords pour que des processus de préservation comme la copie puissent se dérouler ;
- les mesures initiales pour maintenir l'accès peuvent inclure la conservation de tous les logiciels requis pour l'accès ainsi que le matériel spécialisé. Ce ne sera pas une stratégie efficace à long terme mais elle peut se révéler indispensable à court terme ;
- il peut être nécessaire d'évaluer les matériaux numériques, de décider de la durée pendant laquelle ils doivent être conservés et par qui, conformément à une politique approuvée, tel un service de mise au rebut d'archives.

## POUR LES PROGRAMMES DE PRESERVATION AYANT PEU DE RESSOURCES

### **13.14 Réduire les charges**

Les programmes de préservation numérique ayant peu de ressources risquent de découvrir qu'ils sont incapables de disposer de ressources suffisantes pour travailler avec des producteurs. Cependant, ils peuvent parfois faire un investissement valable qui réduira les coûts futurs en entreprenant des actions limitées et ciblées visant à influencer sur le matériau qu'ils doivent gérer. Par exemple :

- s'engager avec seulement un ou deux producteurs pour sonder ce qui est réalisable peut permettre de découvrir quelques mesures simples sur lesquelles il est possible de se mettre d'accord ;
- limiter la gamme des matériaux sélectionnés à quelques formats bien normalisés peut faciliter la fourniture de spécifications que les producteurs suivront sans avoir besoin d'une contribution individuelle ;
- utiliser des directives existantes préparées pour d'autres programmes peut atteindre le même objectif, pour autant que les directives sont appropriées. (Un bon nombre de ces directives sont disponibles auprès d'organismes comme la Library of Congress, des archives de données qui font partie du Arts and Humanities Data Service (Royaume-Uni). Divers organismes ont aussi négocié des accords de licence avec des éditeurs

commerciaux, par exemple, qui peuvent fournir de bons modèles de discussion avec les producteurs locaux.)

### 13.15 Répartir les charges

Les programmes de préservation peuvent aussi trouver des partenaires qui souhaitent partager les charges de la liaison avec les producteurs :

- ils peuvent trouver une institution partenaire avec un programme doté de meilleures ressources qui a déjà établi de bons dispositifs de travail avec un groupe de producteurs. Dans le cadre d'un accord de développement, les producteurs peuvent souhaiter y inclure d'autres partenaires dans la mesure où il existe des garanties appropriées à leurs intérêts ;
- un certain nombre de programmes plus petits qui travaillent dans la même région auront peut-être la possibilité de former un consortium en vue de négocier des arrangements avec les producteurs pour le compte de tous.

### ETUDES DE CAS

Le tableau 13.3 présente quelques scénarios possibles dans différents environnements.

### REFERENCES - pour en savoir plus

#### Renvois internes

*Les problèmes de droits*, voir aussi Gérer les droits : chapitre 15

#### Renvois externes (tous les liens ont été consultés à la date de mars 2003)

- Arts and Humanities Data Service (sans date). (*Digitisation: a Project Planning Checklist*. <http://ahds.ac.uk/checklist.htm>)
- Canadian Heritage Information Network (CHIN) (2002). *Creating and Managing Digital Content*. [http://www.chin.gc.ca/english/digital\\_content/index.html](http://www.chin.gc.ca/english/digital_content/index.html)
- Cornell University Library (2003) *Moving Theory into Practice: Digital Imaging Tutorial*. <http://www.library.cornell.edu/preservation/tutorial/contents.html>
- Digital Library Federation (2002). *Benchmarks for Digital Reproductions of Monographs and Serials*. <http://www.diglib.org/standards/bmarkfin.htm-bench>
- Humanities Advanced Technology and Information Institute (HATII), Université de Glasgow et National Initiative for a Networked Cultural Heritage (NINCH) (2002). *The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials*. <http://www.nyu.edu/its/humanities/ninchguide/>
- Institute of Museum and Library Services (IMLS) (2001). *A Framework of Guidance for Building Good Digital Collections*. <http://www.imls.gov/pubs/forumframework.htm>
- Library of Congress (2002). *Building Digital Collections: Technical Information and Background Papers [for American Memory Programme]*. <http://lcweb2.loc.gov/ammem/ftpfiles.html>

- MATRIX: (The Center for Humane Arts, Letters and Social Sciences Online, Université d'Etat du Michigan) (sans date). Working Paper on Digitizing Audio for the Nation Gallery of the Spoken Word and the African Online Digital Library. <http://africandl.org/bestprac/audio/audio/html>
- National Library of Australia (2002). Safeguarding Australia's web Resources: guidelines for creators and publishers. <http://www.nla.gov.au/guidelines/webresources.html>
- Natural Environment Research Council (NERC) (UK) (2002). NERC Data Policy Handbook. <http://www.nerc.ac.uk/data/documents/datahandbook.pdf>
- Pockley, Simon (1998). Cinemedia: Good Practice Book. <http://www.acmi.net.au/FOD/DuckDigital/GoodP.html>
- Rowe, J. (2002). "Developing a 3D Digital Library for Spatial Data: Issues Identified and Description of Prototype" in RLG DigiNews, 6(5). <http://www.rlg.org/preserv/diginews6-5.html-feature1>
- Townsend, S., Chappell, C., Struijve, O. (1999). Digitizing History: a Guide to Creating Digital Resources from Historical Documents. (Un guide AHDS de bonne pratique.) [http://hds.essex.ac.uk/g2gp/digitising\\_history/index.asp](http://hds.essex.ac.uk/g2gp/digitising_history/index.asp)

---

## **Bibliothèque avec droits de préservation**

---

### Niveau probable d'influence ou de contrôle

---

Faible : il peut être difficile d'identifier les producteurs ou de les contacter

Il peut y avoir un ensemble de producteurs extrêmement varié (par exemple les éditeurs du Web) des problèmes avec les droits commerciaux

L'incapacité de spécifier les formats à utiliser

---

### *Stratégies susceptibles d'exercer une influence*

---

Identifier les producteurs représentatifs et travailler avec eux

Essayer d'établir un code de pratique

Un programme d'éducation, des séminaires, des directives

Mettre l'accent sur les avantages pour les producteurs

Peut réclamer une négociation individuelle

Peut nécessiter de chercher une loi de dépôt légal

---

### **Bibliothèque avec accès sous licence**

---

---

#### Niveau probable d'influence ou de contrôle

---

Bon contrôle sur les formats mais éventuellement contrôle réduit sur la préservation

---

#### *Stratégies pouvant exercer une influence*

---

Peut nécessiter des négociations prolongées pour assurer l'accessibilité continue

---

### **Archives institutionnelles**

---

---

#### Niveau probable d'influence ou de contrôle

---

Potentiellement bon si capable d'établir des spécifications et des procédures normatives

Peut avoir des sanctions légales ou organisationnelles pour faire respecter la conformité

---

#### *Stratégies pouvant exercer une influence*

---

Programme d'éducation et soutien technique pour encourager la conformité

Influencer la spécification, la conception et l'acquisition de systèmes et de pratiques de conservation d'enregistrement

---

### **Archives spécialisées audiovidéo**

---

---

#### Niveau probable d'influence ou de contrôle

---

Peut être bon pour les déposants commandités mais faible pour les autres

Il peut y avoir un usage de formats et normes différents

---

*Stratégies pouvant exercer une influence*

---

Etablir des formats normalisés pour qu'ils soient acceptés dans le programme

Elaborer un code de pratique avec les groupes professionnels ou les dirigeants professionnels

Développer des relations étroites avec le groupe des producteurs

---

**Archives de données**

---

Niveau probable d'influence ou de contrôle

---

Susceptible d'être bon si son rôle et sa crédibilité sont reconnus par les groupes de producteurs et d'utilisateurs

---

*Stratégies pouvant exercer une influence*

---

Proclamer ses besoins

Aider les producteurs à concevoir leurs projets

Encourager le dépôt de l'information contextuelle

---

**Archives de projet communautaire**

---

Niveau probable d'influence ou de contrôle

---

Susceptible d'être bon si l'implication est précoce dans la planification du projet mais limité si elle est tardive

---

*Stratégies pouvant exercer une influence*

---

Travailler en partenariat étroit avec les producteurs

Fournir les outils qui permettent aux participants de la communauté d'être compatibles

Intégrer un programme de préservation dans les objectifs du projet communautaire

---

**Tableau 13.3 - Quelques possibilités de travail avec les producteurs dans divers secteurs**

## Chapitre 14. Prendre le contrôle : transfert et métadonnées

### INTRODUCTION

#### 14.1 Les objectifs

Ce chapitre propose à la fois des conseils de gestion et des conseils techniques sur des points relatifs au contrôle du patrimoine numérique par les programmes de préservation.

#### 14.2 En bref

Contrôler ce qui arrive aux matériaux numériques est une mesure de préservation essentielle. Dans la plupart des cas, cela nécessite le transfert des données et de la documentation aux soins d'un programme de préservation, où on leur attribue une identification unique, et où elles sont décrites par le moyen de métadonnées de diverses sortes. Les métadonnées permettent de trouver les matériaux numériques, et d'une manière essentielle du point de vue de la préservation, de les gérer et de les représenter avec exactitude. Même si les normes de préservation des métadonnées continuent d'évoluer, les programmes doivent décrire les caractéristiques techniques, la provenance et les objectifs de préservation des matériaux numériques qui leur sont confiés.

### LES PRINCIPAUX PROBLEMES DE GESTION

#### 14.3 Transférer les données dans un lieu sûr

Ces Directives recommandent le transfert des matériaux du patrimoine numérique d'un environnement de fonctionnement dans un lieu sûr pour éviter les risques de détérioration ou de perte liés à une utilisation quotidienne des fichiers numériques. Dans la plupart des cas, cela nécessite le transfert des données aux soins d'un organisme de préservation responsable.

Le processus de transfert lui-même n'est pas sans risques car les données à cette occasion peuvent être perdues, modifiées, mal identifiées, ou isolées du contexte qui leur donnent leur sens.

#### 14.4 Les problèmes de droits

Du fait que les producteurs de matériaux numériques ont en général des droits sur les matériaux qu'ils produisent, le transfert soulève un certain nombre de problèmes de droits juridiques et moraux.

#### 14.5 Imposer un contrôle

Une fois que les matériaux numériques ont été transférés, ils doivent être contrôlés et organisés d'une manière effective et efficace. Il faut en général pour cela que les matériaux puissent être facilement localisés, accessibles, utilisés, gérés et préservés, conformément aux autorisations.

## DES PRINCIPES POUR RELEVER CES DEFIS

### 14.6 Construire sur des pratiques établies

Le transfert et le contrôle sont des pratiques établies depuis longtemps dans la gestion du patrimoine non numérique. On doit modifier leurs processus quand on les applique à des matériaux numériques.

Une base juridique adéquate pour le transfert est nécessaire. Elle doit traiter les problèmes touchant la facilité avec laquelle les matériaux peuvent être réutilisés, ainsi que la nécessité de copier les données pour les préserver.

Le transfert doit s'effectuer sans perte de données, en utilisant souvent des méthodes très différentes de celles employées pour le transfert des matériaux non numériques.

Le transfert de la documentation accompagnante est particulièrement important pour les données numérique qui ne sont pas compréhensibles sans elle.

### 14.7 Deux méthodes pour effectuer des transferts

La plupart des stratégies de transfert sont des variantes de deux concepts fondamentaux : les producteurs *poussent* les matériaux numériques dans le programme de préservation, ou le programme de préservation *tire* les matériaux de chez le producteur.

Les responsables de programme doivent décider de la méthode la plus appropriée pour le transfert des matériaux et pour le déroulement du travail de chacune des parties impliquées.

### 14.8 Contrôler les formats et les normes

De nombreux programmes imposent des contrôles au moment du transfert sur formats du matériau qu'ils reçoivent. Le but en est de simplifier la préservation en réduisant les variantes qui doivent être gérées en stockant le matériau et en le gardant accessible. Tous les programmes ne sont pas à même de réduire les formats qu'ils acceptent, mais ils doivent s'efforcer de vérifier que les formats ont été utilisés de façon standard.

### 14.9 Contrôler le matériau par l'identification

Il faut donner aux fichiers numériques des identifiants de fichier appropriés pour qu'on puisse les retrouver. Chaque fichier faisant partie d'un système de stockage doit être identifié par un nom de fichier unique de sorte qu'on ne puisse le confondre avec un autre.

Il est également capital pour les programmes de préservation de garantir que les matériaux qu'ils conservent peuvent être trouvés avec certitude, quel que soit leur emplacement. Le Universal Resource Locator (URL) utilisé pour identifier les ressources provenant de la Toile, par exemple, ne permet pas aux utilisateurs de trouver un matériau une fois qu'il a été déplacé. Résoudre ce problème nécessite une forme d'*identification permanente*, construite autour d'un identifiant et d'un moyen de *résolution* ou de connexion au fichier dans son emplacement actuel. Un certain nombre de systèmes sont proposés ou en place comme le Digital Object Identifier (DOI) qu'utilisent les éditeurs, et divers systèmes consultés par les bibliothèques et les archives, mais à ce jour aucun n'a été universellement accepté.

#### 14.10 Contrôler le matériau par la description

Les programmes de préservation utilisent des métadonnées - des informations structurées sur les ressources des données - pour décrire les matériaux numériques qui leur sont confiés. Il y a au moins trois raisons essentielles pour décrire en détail les matériaux du patrimoine numérique :

- pour qu'ils soient trouvés, évalués, rendus disponibles et compris. Cette nécessité a conduit au développement des métadonnées de *découverte de ressources* qui vont des simples listes de noms de fichier aux descriptions approfondies qui encapsulent une riche documentation contextuelle. Les systèmes de métadonnées de "*découverte de ressources*" tels que le Dublin Core, le MARC, les normes de description d'archivage et les catalogues de musée, sont des outils importants à prendre en compte et à utiliser comme étant adaptés à leurs besoins par les programmes de préservation ;
- pour que l'on puisse gérer le déroulement des opérations. Les programmes de préservation engendrent de grandes quantités d'informations sur la façon dont le matériau est créé, transféré et utilisé ; sur les droits et sur qui est autorisé à faire quoi ; et d'autres processus de gestion. Le US NISO *Data Dictionary-Technical Metadata for digital still images*, publié comme projet de norme en 2002, est un exemple d'un ensemble très étendu de "*métadonnées de découvertes de ressources*" disponible sur l'Internet à : [http://www.niso.org/standards/resources/Z39\\_87\\_trial\\_use.pdf](http://www.niso.org/standards/resources/Z39_87_trial_use.pdf) ;
- pour que les programmes de préservation puissent comprendre la manière de représenter les matériaux numériques lorsqu'on a besoin d'y accéder. Les *métadonnées de préservation* décrivent les moyens de fournir l'accès, en même temps que les éléments des métadonnées de gestion de ressources nécessaires pour la gestion des processus de préservation. Elles sont essentielles pour tout programme de préservation ; une conception et une gestion attentives de celles-ci sont particulièrement importantes pour les grandes collections qui doivent être exploitées sous la forme la plus automatisée possible.

#### 14.11 Les métadonnées comme ressources d'informations

Les métadonnées sont elles-mêmes des ressources d'information qu'il faut gérer et préserver, en même temps que le matériau qu'elles décrivent.

#### 14.12 Une approche normalisée des métadonnées

Des systèmes de métadonnées élaborés individuellement peuvent réussir à décrire des collections de matériaux numériques, mais un nombre croissant de raisons plaident en faveur d'une approche normalisée en accord avec d'autres systèmes largement adoptés :

- pour réduire les coûts considérables que représente l'élaboration de systèmes individuels ;
- pour profiter des outils logiciels disponibles qui reconnaissent et enregistrent automatiquement les éléments normalisés des métadonnées à partir des matériaux numériques, réduisant considérablement le coût de la saisie des métadonnées ;
- pour permettre aux programmes de préservation de partager l'information, en rendant leurs collections visibles et consultables pour un public plus large ;
- pour permettre aux matériaux de la collecte d'être déplacés d'un gisement de données à un autre sans avoir besoin d'une réécriture en bloc des métadonnées ;

- pour encourager la normalisation des processus de préservation qui sont décrits et contrôlés par les métadonnées.

Indépendamment des métadonnées de préservation, qui sont plus amplement traitées ci-dessous, il n'entre pas dans le cadre de ces Directives de fournir des informations supplémentaires sur les métadonnées. Par principe, les responsables de programmes de préservation doivent prendre connaissance des systèmes de métadonnées normalisés qui sont largement utilisés dans leur domaine d'intérêt, et adopter ceux qui répondront le mieux à leurs besoins. Ils doivent aussi surveiller l'évolution des normes des métadonnées à laquelle travaillent divers groupes internationaux qui s'intéressent à la gestion des ressources numériques.

## PROBLEMES TECHNIQUES ET PRATIQUES

### 14.13 Lancer le transfert des données

- La reproductibilité des matériaux numériques signifie que le transfert ne nécessite plus de déplacer le matériau d'un site à un autre. On peut transférer des copies parfaitement authentiques pour la préservation tandis que les copies "en direct" restent chez le créateur.
- Le calendrier du transfert peut être d'une importance cruciale. Même si le matériau a été sélectionné pour la préservation, la sélection en soi n'a aucun pouvoir pour ralentir des processus comme la détérioration des supports ou l'obsolescence de la technologie. Le transfert doit s'effectuer assez rapidement pour prévenir ces menaces.
- Le processus de transfert peut exiger d'inclure, outre les fichiers sélectionnés :
  - le transfert de la documentation (y compris le conditionnement pour les supports de format physiques qui sont édités comme les CD et les disquettes), les règles de données et les informations sur la provenance et le contexte d'origine ;
  - le transfert des métadonnées existantes ;
  - les informations sur les droits, y compris les accords de licence ;
  - les informations sur les moyens de fournir l'accès, et peut-être les moyens eux-mêmes - tout logiciel particulier et même le matériel - dont on a besoin pour y accéder actuellement.

### 14.14 Spécifier les supports et les formats de fichier

Il n'y a pas de méthode normalisée pour effectuer le transfert physique de matériaux numériques. Les données peuvent être transférées sur une gamme étendue de supports physiques tels que, sous des formes diverses, de disquettes, CD, bandes, cartouches et lecteurs de disquettes ; ou par des réseaux de communication utilisant des moyens comme les connexions du courriel, le File Transfer Protocol (FTP) ("protocole de transport de fichiers"), et le transfert à partir des sites Web. Le choix des supports de transfert dépend des besoins des tiers concernés.

Quels que soient les moyens choisis, les données doivent rester en sécurité. Certains environnements de transfert peuvent présenter des risques particuliers pour des supports spécifiques ; par exemple, les supports physiques sont facilement perdus ou volés, tandis que des réseaux de communication se révéleront non fiables et il peut être plus sûr de remettre en main propre un support physique.

Le support de transfert doit permettre de charger et de récupérer les données. Lorsque l'expéditeur et le récepteur utilisent tous les deux les mêmes technologies, le transfert sera relativement facile. Lorsque les technologies sont mal assorties, l'un des deux devra supporter le coût occasionné par l'utilisation de technologies différentes.

Les programmes peuvent avoir des dispositifs pour manipuler une large gamme de supports, ou disposer de dispositifs plus limités. Les supports physiques nécessitent un équipement informatique spécifique que les organismes de préservation peuvent n'être pas en mesure de fournir. En pareil cas, ils devront décider s'il est raisonnable d'exiger des transferts via les supports spécifiques qu'ils peuvent réaliser, ou investir dans des dispositifs pour manipuler une gamme plus large de supports.

Ci-dessous, le tableau 14.1 présente quelques considérations pour choisir des supports de transfert.

---

Si les données doivent rester sur le support de transfert pour un stockage à moyen terme,

---

éviter les supports de court terme comme les disquettes ou les bandes audionumériques.

---

Si les données doivent être immédiatement chargées sur un autre support pour le stockage,

---

un support de court terme peut convenir pour le stockage

---

Si les coûts de logement pour une large gamme de supports sont prohibitifs,

---

cibler une gamme plus restreinte de supports

---

Si le flux de travail est construit autour d'un support spécifique,

---

cibler des supports qui soient adaptés au flux de travail, ou cibler des supports sur lesquels il sera facile pour les producteurs de fournir et d'ajuster le flux de travail.

---

**Tableau 14.1 - Facteurs de décision dans le choix des supports de transferts**

### **14.15 Les stratégies de transfert**

Les transferts de données impliquent généralement les programmes de préservation soit parce qu'ils reçoivent des fichiers du producteur (technique de transfert à *flux poussé*), soit qu'ils retirent des fichiers du site du producteur (technique de transfert à *flux tiré*).

Il existe de nombreuses techniques à *flux poussé* ou de dépôt, comme envoyer des fichiers chargés sur un support physique par courrier ou par coursier, joindre des fichiers à des messages électroniques ; ou les transmettre par le FTP directement au serveur du programme de préservation. Les techniques à *flux poussé* ont de nombreux avantages : elles permettent au producteur de

déposer plus de versions facilement préservées de son travail qu'il n'y en a à la disposition du public, et elles donnent au producteur une plus grande possibilité d'influer sur la sélection de ce qui sera préservé.

En revanche, les programmes de préservation qui s'appuient sur le dépôt peuvent découvrir que les transferts dépendent de facteurs de production qui échappent à leur contrôle, notamment les changements de personnel, les modifications de priorités, ou des niveaux d'intérêt en déclin, tous conduisant à des transferts incohérents

Les techniques à *flux tiré* accordent un pouvoir de contrôle plus important au programme de préservation en ce qui concerne le calendrier et le contenu des transferts. Certains producteurs considèrent cela comme une violation de leurs droits : soit ils barrent le logiciel utilisé pour copier leurs fichiers, soit ils réclament des accords de droits, aussi le contrôle offert par les techniques à *flux tiré* n'est-il pas absolu. (D'un autre côté, de nombreux producteurs sont heureux de voir leur matériau saisi, préservé et rendu disponible sans que cela leur coûte rien.)

La *collecte* ou la *récolte* automatisée du matériau issu des sites du producteur peut se faire par les réseaux de communication. En utilisant des logiciels programmés pour explorer le réseau en vue de trouver des fichiers qui correspondent aux critères, les programmes de préservation peuvent copier et télécharger des fichiers sur leurs propres systèmes informatiques. Cette méthode est largement utilisée par les moteurs de recherche de l'Internet et par la plupart des programmes de préservation qui saisissent le matériau mis en réseau. Divers programmes de logiciels d'indexation et de "recherche et récupération" sont disponibles, avec des capacités variables pour définir ce qui doit être ou non récupéré.

La collecte peut être un moyen très efficace de saisie de données, mais elle peut aussi poser des problèmes. Certains dossiers risquent d'être invisibles pour le logiciel : ils ne peuvent être accessibles que par une interface d'exploitation qui communique avec des données sous-jacentes. De nombreux producteurs stockent également des versions de leur travail de meilleure qualité, telles que des fichiers d'images et audio, en les séparant des versions dérivées qui sont adaptées à la livraison des réseaux : la collecte rate les versions qui doivent être préservées et saisit les versions prévues seulement pour un accès à court terme.

On résout souvent ces dilemmes grâce à des accords mixtes aux termes desquels les producteurs acceptent de mettre une version appropriée de leur travail là où le programme de préservation peut la collecter.

#### **14.16 Le contrôle de la qualité**

Mis à part les problèmes de transfert, les programmes de préservation doivent contrôler le matériau dès réception pour confirmer que tous les fichiers demandés ont été reçus, qu'ils fonctionnent comme prévu, et que les métadonnées et autre documentation sont conformes.

#### **14.17 L'identification des fichiers**

Les objets numériques peuvent avoir un certain nombre d'identifiants utilisés pour le contrôle sur place, pour l'identification de tout le système, et pour l'accès global (exactement comme un livre sur l'étagère d'une bibliothèque peut être identifié par son titre, un numéro de classification, un emplacement d'étagère, un numéro matricule, un numéro d'enregistrement dans la base de données du catalogue, un numéro ISBN, etc.).

Une identification permanente est nécessaire pour que l'on puisse trouver les éléments même s'ils sont déplacés dans un système de stockage. Les liens incorporés dans les objets ne continueront à fonctionner que s'ils sont reliés à des identifiants permanents.

Les méthodes de remplacement comprennent notamment :

- à l'intérieur d'un petit système, la garantie que tous les utilisateurs soient informés des modifications d'emplacement ;
- des messages de réacheminement automatique qui conduisent les utilisateurs au nouvel emplacement ;
- une gestion du stockage des dossiers pour réduire au minimum les déplacements de fichiers ;
- l'utilisation d'un système d'identifiant permanent (IP), ce qui implique un nom de fichier unique et un abonnement à un service de serveur qui enregistre les IP et leur emplacement actuel. (Les programmes doivent noter que les systèmes IP existants sont, soit encore grandement sous-développés, soit disponibles, mais y participer coûte cher.)

#### **14.18 Prendre soin des métadonnées**

On a déjà noté qu'il fallait non seulement enregistrer les métadonnées, mais aussi en prendre soin. D'où un certain nombre de dispositions :

- Structurer. Organiser les métadonnées dans une structure de document normalisée comme un modèle XML facilitera la préservation.
- Relier. Les liaisons entre les enregistrements de métadonnées et les objets numériques qu'ils décrivent doivent être maintenues. Le lieu de stockage des métadonnées le plus favorable à cette opération est l'objet de nombreuses discussions. Alors que certaines métadonnées doivent être reliées pour que les outils logiciels puissent automatiquement traiter le matériau, doit-on stocker la totalité des enregistrements de métadonnées séparément, ou bien les relier, ou encore les intégrer aux objets qu'elles décrivent ? Ce choix pose problème. Un stockage séparé permet d'accéder aux métadonnées et de les mettre à jour sans avoir besoin de sortir les objets numériques reliés du stockage - ce qui est un avantage important. D'un autre côté, de nombreux responsables de programme s'inquiètent de la possibilité que la liaison essentielle entre l'objet et l'enregistrement soit interrompue sur de grandes périodes de temps. Les responsables doivent évaluer les risques encourus et décider quelle est la méthode la plus adéquate.
- Contrôler la qualité. Garantir la fiabilité des enregistrements de métadonnées est une priorité essentielle. Des mesures de contrôle de la qualité sont nécessaires chaque fois que les métadonnées sont créées ou modifiées.
- Protéger. L'intégrité des enregistrements de métadonnées doit être assurée : elles réclament la même attention de préservation que les objets qu'elles décrivent.

### **14.19 Préparer le paquet d'archivage pour le stockage**

Une fois que le matériau numérique a été transféré et le travail de contrôle et de description effectué, il doit être préparé pour entrer dans un système de stockage. Cela doit être fait en s'assurant que les diverses parties du paquet d'information, y compris le contenu et les métadonnées, sont reliées et qu'un train de données est créé qui soit stockable en toute sécurité sur le support de stockage en usage et qui puisse être trouvé par les programmes de recherche de fichier appropriés.

Le paquet est alors envoyé pour sauvegarde au stockage.

Avant de mettre l'objet numérique en stockage en tant qu'original de préservation, de nombreux programmes créent des copies supplémentaires pour au moins deux excellentes raisons :

- afin d'avoir une copie disponible prête à l'emploi sans avoir besoin de sortir l'original de préservation du stockage. Les copies d'utilisation sont souvent optimisées pour accéder aux communications actuellement disponibles et aux technologies d'affichage (comme la basse résolution, les versions compressées de fichiers d'image qui peuvent être livrées beaucoup plus rapidement sur le Réseau). Il n'est généralement pas nécessaire de préserver les copies d'accès dérivées au fur et à mesure des changements de technologie et, souvent, elles n'ont pas d'enregistrements de métadonnées de préservation détaillés ;
- afin de stocker les objets sur plus d'un format, amorçant ainsi des stratégies de remplacement pour fournir un accès dans l'avenir. Comme on l'a vu au chapitre 16, c'est une bonne pratique que de conserver des copies des objets numériques dans leurs formats originaux, indépendamment de la nécessité de créer de nouveaux formats en tant qu'original de préservation ou copie d'accès courant.

Bien entendu, toutes les versions parallèles doivent être gérées comme des objets numériques distincts mais connexes.

## **CONSIDERATIONS PARTICULIERES**

### **14.20 Les métadonnées de préservation**

Les métadonnées de préservation sont des informations structurées relatives à un objet numérique qui :

- identifient le matériau dont un programme de préservation a la responsabilité ;
- communiquent ce qui est nécessaire pour conserver et protéger les données ;
- communiquent ce qui est nécessaire pour représenter l'objet voulu (ou ses éléments essentiels tels qu'ils ont été définis) à un utilisateur lorsqu'il en a besoin, indépendamment des changements survenus dans le stockage et les technologies d'accès ;
- enregistrent l'historique et les conséquences de ce qui arrive à l'objet ;
- documentent l'identité et l'intégrité de l'objet en tant que fondement de son authenticité ;
- permettent à l'utilisateur et au programme de préservation de comprendre le contexte de l'objet dans le stockage et dans l'utilisation.

Les dispositifs d'enregistrement des métadonnées de préservation doivent accepter le fait que le même contenu fondamental (ou objet conceptuel) peut exister sous diverses formes au cours de son existence. Certaines de ces formes coexisteront en tant qu'objets numériques alors que d'autres se succéderont dans une suite de générations qui sont séparées ou qui se chevauchent. Cela transparaît dans certains programmes qui créent un enregistrement pour une seule et unique version identifiée comme "original de préservation", en le documentant par ses variantes et ses modifications en tant que partie intégrante de l'histoire de cet objet.

D'autres programmes créent un enregistrement pour chacune des formes qui nécessitent une action de préservation, en veillant à ce que les rapports entre les formes soient explicites dans les enregistrements des métadonnées.

Les informations nécessaires pour la préservation des métadonnées sont souvent divisées en deux catégories (conformément au Reference Model for an Open Archival Information System ou OAIS mentionné au chapitre 8) :

- *les informations de contenu*, principalement constituées de détails sur la nature technique de l'objet qui précise au système comment représenter les données en tant que types et formats de données spécifiques. Au fur et à mesure que les technologies d'accès changent, ces métadonnées de représentation doivent aussi être actualisées ;
- *les informations décrivant la préservation*, constituées d'autres informations nécessaires pour une gestion à long terme et une utilisation de l'objet, comprennent les identifiants et les détails bibliographiques, des informations sur la propriété et les droits, la provenance, l'histoire, le contexte, y compris les rapports avec les autres objets, et les informations quant à la validation.

Bien entendu, certaines de ces métadonnées peuvent se rapporter à d'autres objets d'information comme les outils logiciels et les spécifications de format qui doivent aussi être gérés. La nature interdépendante des matériaux numériques signifie que les programmes doivent souvent gérer des réseaux d'objets connectés et leurs métadonnées.

Il n'existe toujours pas de normes reconnues d'utilisation universelle pour les schémas des métadonnées de préservation, aussi les programmes doivent-ils choisir entre adopter (et probablement adapter) l'un des modèles utilisés par d'autres, ou concevoir leur propre schéma (soit comme solution globale, soit comme une solution intermédiaire minimale jusqu'à ce qu'une norme apparaisse).

De nombreux services nationaux ont publié des spécifications de métadonnées pour des systèmes de conservation d'enregistrements qui incluent les besoins de la préservation. Dans le domaine de la bibliothèque, un groupe de travail international réuni par l'OCLC et le Research Libraries Group (RLG) a publié un cadre préconisé de métadonnées de préservation au milieu de l'année 2002 (disponible sur le réseau à <[http://www.oclc.org/research/pmwg/pm\\_framework.pdf](http://www.oclc.org/research/pmwg/pm_framework.pdf)>). Leur rapport est un bon point de départ pour explorer les métadonnées qui seront nécessaires.

La Bibliothèque nationale de Nouvelle-Zélande a réalisé un projet intéressant qui tente d'adapter le travail de l'OCLC et du RLG à un programme particulier et à ses circonstances propres (disponible sur le Réseau à <[http://www.natlib.govt.nz/files/4initiatives\\_metaschema.pdf](http://www.natlib.govt.nz/files/4initiatives_metaschema.pdf)>). Ce schéma propose en résumé les éléments suivants :

Décrire un objet numérique

Nom de l'objet

Identifiants locaux

Identifiant permanent global

Emplacement du fichier dans le système de stockage

Date à laquelle a été créé l'original de préservation

Composition technique dominante (n° des fichiers de chaque type de codage MIME)

Type de structure (par exemple le texte, l'image)

Matériel informatique nécessaire pour que l'objet fonctionne

Logiciel nécessaire pour que l'objet fonctionne

Instructions particulières d'installation

Inhibiteurs et facilitateurs d'accès intégrés

Anomalies intégrées

Clés d'authentification et de validation

Date de la création

Type/format de codage MIME (par exemple image/mode point)

Version

Fichier clé qui fournit l'accès

Caractéristiques des types de fichier spécifique :

(par exemple pour les fichiers d'image : la résolution, les dimensions, la résolution tonale, la place des couleurs, la gestion des couleurs, la table des couleurs, l'orientation, la compression)

(par exemple pour les fichiers de texte : la compression, l'alphabet, la fonction DTD associée pour le texte structuré, les divisions de structure)

(par exemple pour l'audio : la résolution, la durée, le débit binaire, la compression, l'encapsulation, le numéro de piste et le type)

(par exemple pour la vidéo : les dimensions de l'image, la durée, le nombre d'images à la seconde, la compression, la structure de codage, le son)

(par exemple pour les ensembles de données : les éléments communs ci-dessus seulement)

(par exemple pour les fichiers exécutables : les éléments communs ci-dessus seulement)

Décrire l'actualisation des métadonnées

Agent modifiant les métadonnées

Date de la modification

Domaine modifié

Qui a créé les métadonnées et quand ?

Décrire tous les processus appliqués à un objet (y compris la création)

Nom du processus

Objectif

Agent qui a exécuté le processus

Agent qui a homologué le processus et quand

Matériel utilisé

Logiciel utilisé

Opérations ayant fait partie du processus

Résultats

Normes ou spécifications utilisées

Date de son achèvement

Décrire les caractéristiques techniques de tous les fichiers à l'intérieur de l'objet

Identifiants de fichier spécifique

Rapport avec les autres fichiers de base

Taille de fichier

**POUR LES PROGRAMMES DE PRESERVATION DISPOSANT  
DE PEU DE RESSOURCES**

**14.21 Le transfert**

Les programmes de préservation disposant de peu de ressources auront besoin de rechercher des moyens de réduire les coûts de transfert :

- les techniques du *flux poussé* peuvent réclamer un investissement moindre de la part du programme de préservation et faire passer la majeure partie du coût sur le producteur. Toutefois, sans accords sur les supports, les formats et le contrôle de qualité qu'utiliseront les producteurs lors du transfert du matériel, des économies à court terme risquent d'entraîner des coûts de préservation plus élevés à long terme ;

- une limitation bien pensée du registre des supports et des formats admis par le programme permettra de faire des économies ;
- les programmes peuvent être en mesure de stocker le matériau transféré sur leurs supports de transfert s'ils ont choisi des supports suffisamment stables, et s'il peuvent faire par sécurité des copies de sauvegarde.

Certains groupes qui n'ont pas accès à des organismes de préservation distincts devront suivre un modèle de "non-transfert", en installant les meilleurs dispositifs de préservation possibles dans un environnement d'exploitation. Même dans ces circonstances, nombre de ces mêmes principes s'appliquent : l'accessibilité continue est plus probable avec une sorte de transfert interne sur "un centre d'archivage de sauvegarde", même modeste, où des fichiers peuvent être gérés en dehors des risques normaux de l'utilisation opérationnelle. Les fichiers devront toujours être suffisamment bien décrits et protégés pour permettre un transfert ultérieur sur un programme de préservation plus sûr.

#### **14.22 Les métadonnées**

Les coûts d'enregistrement des métadonnées peuvent représenter une part importante des coûts de préservation globaux. On peut faire des économies soit en réduisant la quantité des informations enregistrées au minimum (et en acceptant qu'à la fois l'accès et la préservation soient rendus plus difficiles), soit en investissant dans un logiciel qui saisira automatiquement les métadonnées (ce qui deviendra plus facile au fur et à mesure que les normes des métadonnées se développent).

En choisissant un ensemble de métadonnées minimal, les responsables de programme jugeront peut-être utile de prendre en considération les besoins futurs des utilisateurs afin de trouver le matériau, et les questions qui exigeront des réponses en entreprenant toute action de préservation que l'on peut prévoir.

### **ETUDES DE CAS**

#### **14.23 Etude de cas 1**

Une bibliothèque nationale qui collecte des publications numériques en ligne réalise des transferts en recueillant des fichiers sur des sites d'éditeurs, en utilisant des recherches programmables, en copiant et en téléchargeant des logiciels comme l'armoire HTTP, conformément aux accord négociés avec chacun des propriétaires de site. Le processus de collecte exige du personnel pour chercher des sites potentiels qui puissent correspondre aux directives de sélection de la bibliothèque, pour décider ce qui doit être saisi et jusqu'où les liens sur le site doivent être suivis (la politique de sélection conseille de saisir les documents connectés sur le même site, mais de ne pas suivre d'autres liens). Lorsque les fichiers sont téléchargés par le logiciel, le personnel vérifie que la totalité du matériau désiré a été téléchargé et que tous les fichiers fonctionnent. Un enregistrement des métadonnées est créé en utilisant un mélange de données produites par le logiciel et de données rentrées manuellement. Une page d'entrée est créée pour chaque titre saisi, en utilisant un modèle engendré par le système, de sorte que les utilisateurs puissent comprendre ce qu'ils vont obtenir et comment cela est apparenté à la fois au site Web de l'éditeur et à d'autres matériaux saisis dans le centre d'archivage. Lorsqu'il est achevé, l'enregistrement des métadonnées qui inclut un lien envoyant aux objets saisis, est envoyé pour sauvegarde au gisement de données des métadonnées, et les objets saisis sont envoyés pour sauvegarde au système de stockage de grande capacité de gisement de données.

## 14.24 Etude de cas 2

Un petit centre d'archivage d'ethnomusicologie reçoit de collectionneurs des enregistrements réalisés sur le terrain sur une bande audionumérique, qui est bon marché et commode pour la collecte mais inadéquate pour le stockage. Après avoir vérifié que le matériau est conforme à la politique de collecte du centre d'archivage, et que la qualité d'enregistrement est adéquate, le personnel accepte le matériau, et rentre manuellement les informations sur le dépôt dans une base de données distincte. Le matériau est porté sur le registre des additions et doté d'un numéro courant dans la collection. Les données qui sont sur la bande sont copiées sur deux jeux de CD-ROM : un comme copie de préservation, un comme sauvegarde. La bande audionumérique originale est rangée sur une étagère comme copie d'accès pour utilisation à court terme, et les copies de CD sont rangées sur une étagère séparément. L'enregistrement des métadonnées est actualisé avec l'emplacement de toutes les copies.

### REFERENCES - pour en savoir plus

#### Références internes

*La liaison avec les producteurs*, voir aussi Travailler avec les producteurs : chapitre 13.

*Les métadonnées et les moyens d'accès*, voir aussi Maintenir l'accessibilité : chapitre 17.

**Références externes** (tous les liens ont été consultés à la date de mars 2003)

#### 1. Le transfert

Le Arts and Humanities Data Service (AHDS) (Royaume-Uni) et ses archives de données affiliées (dans les domaines de la littérature, l'archéologie, les arts visuels, l'histoire et les arts du spectacle) ont publié un certain nombre d'excellents "guides pour les déposants". Ils comportent de bonnes informations techniques sur les formats préférés pour une large gamme de types de matériaux. Ils peuvent servir de bons modèles pour des programmes similaires exploitant des données. Par exemple :

- History Data Service (sans date). *Guidelines for Depositors*. <http://hds.essex.ac.uk/depguide.asp>
- Oxford Text Archive (1999). *Depositing with the OTA: the Depositors Guidelines*. [http://ota.ahds.ac.uk/publications/ID\\_Depositing-Introduction.html](http://ota.ahds.ac.uk/publications/ID_Depositing-Introduction.html)
- Visual Arts Data Service (sans date). *Guidelines for Depositors*. [http://vads.ahds.ac.uk/depositing/depositor\\_guidelines.pdf](http://vads.ahds.ac.uk/depositing/depositor_guidelines.pdf)

#### 2. L'identification permanente

- Corporation for National Research Initiatives (CNRI) (sans date). *The Handle System*. <http://www.handle.net/index.html>
- Dack Diana (2001). *Persistent Identification Systems (Report on a consultancy for the National Library of Australia)*. <Http://www.nla.gov.au/initiatives/persistence/PIcontents.html>

- International DOI Foundation (sans date). *The Digital Object Identifier System*. <http://www.doi.org/>
- Internet Engineering Task Force (IETF) (2001). *Uniform Resource Names (URN)*. <http://www.ietf.org/html.charters/urn-charter.html>
- National Library of Australia (2001). *Managing Web Resources of Persistent Access*. <http://www.nla.gov.au/guidelines/2000/persistence.html>
- PURL Team (sans date). *PURL - Persistent URL Homepage*. <http://purl.oclc.org/>

### 3. Les métadonnées

Les normes et les initiatives adoptées en matière de métadonnées abondent dans les différents domaines de la gestion du patrimoine, avec des extensions et des adaptations pour contenir les matériaux numériques. Voir par exemple :

- Dublin Core Metadata Initiative. <http://dublincore.org/>
- IFLA Universal Bibliographic Control and International MARC Core Programme (UBCIM) (2000). UNIMARC Guidelines n° : Electronic Resources. <http://ifla.org/VI/3p1996-1/guid6.htm>
- International Council on Archives (1999). *General International Standard Archival Description*, 2e édition. [http://www.ica.org/biblio/cds/isad\\_g\\_2e.pdf](http://www.ica.org/biblio/cds/isad_g_2e.pdf)
- Consortium for the Computer Interchange of Museum Information (1999). CIMI Dublin Core Metadata Testbed Project. [http://www.cimi.org/old\\_site/documents/meta\\_webliography.html](http://www.cimi.org/old_site/documents/meta_webliography.html)
- International Association of Sound and Audiovisual Archives (1998). *The IASA Cataloguing Rules*. <http://www.iasa-web.org/icat/>

Quelques sources de métadonnées de préservation :

- Colorado Digitization Project Metadata Workgroup, Audio Taskforce (2002). *Metadata for Digital Audio (draft)*. [http://coloradodigital.coalliance.org/digaudio\\_meta.pdf](http://coloradodigital.coalliance.org/digaudio_meta.pdf)
- National Library of New Zealand (2002). *Metadata standards framework - preservation metadata*. [http://www.natlib.govt.nz/files/4initiatives\\_metaschema.pdf](http://www.natlib.govt.nz/files/4initiatives_metaschema.pdf)
- NISO/AIIM (2002). *Data Dictionary - technical metadata for digital still images*, publié en tant que projet de norme pour mise à l'essai du NISO Z39.87 - 2002. [http://www.niso.org/standards/resources/Z39\\_87\\_trial\\_use.pdf](http://www.niso.org/standards/resources/Z39_87_trial_use.pdf)
- *Preservation metadata and the OAIS Information Model: a metadata framework to support the preservation of digital object: a report by the OCLC/RLG Working Group on Preservation Metadata* (2002). [http://www.oclc.org/research/pmwg/pm\\_framework.pdf](http://www.oclc.org/research/pmwg/pm_framework.pdf)

- Public Record Office (Royaume-Uni) (sans date). *PRONOM* (qui concerne un système de base de données qui stocke et fournit des informations sur les formats de fichier et les logiciels d'application nécessaires pour les ouvrir.) <http://www.pro.gov.uk/about/preservation/digital/pronom.htm>
- La British Library (sans date). *Code of Practice for the Voluntary Deposit of Non-Print Publications*. <http://www.bl.uk/about/policies/codeprac.html>

## Chapitre 15. Gérer les droits

### INTRODUCTION

#### 15.1 Note d'avertissement

*Ces directives ne doivent pas être considérées comme un avis juridique compétent sur les questions de droits.*

#### 15.2 Les objectifs

Ce chapitre entend souligner la responsabilité importante des programmes de préservation qui doivent être avertis de tout ce qui concerne les droits, et faire quelques propositions d'ordre général sur la manière dont il faut aborder ces questions.

#### 15.3 En bref

Les parties prenantes détiennent une série de droits et ont des attentes que les programmes de préservation doivent connaître et, le cas échéant, inclure dans la planification de leur gestion. Nombre de ces droits ont des implications juridiques, y compris les droits de propriété intellectuelle et les droits au respect de la vie privée. Etant donné que les programmes de préservation doivent copier les matériaux numériques pour les préserver, et que la plupart des programmes visent à fournir un certain niveau d'accès, des modes de gestion pratique des droits s'imposent.

### LES PRINCIPALES QUESTIONS DE GESTION

#### 15.4 Le patrimoine numérique et les droits

Les matériaux du patrimoine numérique sont soumis à une série de droits et d'attentes, dont certains ont force de loi. Un grand nombre, comme le copyright, concernent la propriété intellectuelle engagée dans le matériau. Toutefois, il peut y avoir d'autres droits et d'autres attentes qui doivent être pris en compte.

#### 15.5 Une série de droits et d'attentes

La série des droits et des attentes auxquels les programmes de préservation peuvent être confrontés et qu'ils doivent gérer dans un sens précis comprend :

- les droits de propriété intellectuelle des producteurs, y compris le copyright, qui peuvent exister dans les diverses couches liées aux différents aspects des matériaux ; le droit de poser des conditions d'accès et d'utilisation ; et le droit moral du créateur d'être reconnu ;
- les droits légaux de certaines institutions pour collecter, préserver et fournir l'accès ;

- les droits et les attentes touchant le respect de la vie privée, la confidentialité et l'autorisation d'utilisation liés à certains types de matériaux comme les enregistrements organisationnels, les enregistrements d'histoire orale, les données personnelles et les communications privées ;
- les attentes des utilisateurs relatives à l'accès et à l'utilisation ;
- les attentes d'un groupe plus important qui souhaite que le matériau ayant la valeur d'un patrimoine durable soit préservé et rendu accessible avec le régime des droits établi par la loi.

### **15.6 Les droits fondamentaux nécessaires pour les activités de préservation**

La préservation implique de nombreux processus où les questions de droits entrent en ligne de compte. Afin de parvenir à la continuité du patrimoine numérique, les programmes de préservation doivent :

- obtenir et détenir des matériaux, ce qui implique généralement de faire des copies ;
- faire des copies supplémentaires à des fins de préservation ;
- si nécessaire, mettre en dérivation les dispositifs utilisés par les producteurs pour limiter l'accès et empêcher les copies ;
- décider quels matériaux et quels aspects des matériaux doivent être préservés ;
- ajouter les métadonnées ;
- modifier les structures de fichier et les noms de fichier si nécessaire ;
- utiliser tous les moyens présentement disponibles pour préserver l'accessibilité ;
- fournir un accès géré pour les utilisateurs autorisés.

### **15.7 Les défis**

Obtenir des permissions pour mener toutes ces activités n'est pas toujours facile :

- les producteurs et autres propriétaires de droits peuvent ne pas vouloir donner la permission ;
- les droits d'accès et les droits du respect de la vie privée et de la confidentialité sont souvent en opposition ;
- dans un environnement où la création est morcelée ou fait appel à la collaboration de matériaux numériques, il peut être difficile d'identifier tous les propriétaires de droits ou de négocier avec eux ;
- la situation juridique peut être ambiguë, et de nombreuses juridictions sont toujours en train de clarifier les cadres juridiques des droits et la manière dont ils doivent être gérés ;
- en traitant des matériaux mis en réseau à l'échelle du globe, on risque même ne pas très bien savoir quelles sont les dispositions juridiques applicables : celles dans le cadre desquelles ils ont été produits, ou publiés, ou saisis pour la préservation, ou stockés, ou consultés - elles peuvent être différentes pour chaque cas de figure.

Les coûts de la mise en place de bonnes pratiques de gestion des droits peuvent être élevés, en particulier si une négociation individuelle est nécessaire. Mais les coûts suscités par une absence de gestion adéquate des questions de droits peuvent l'être tout autant.

## DES PRINCIPES POUR RELEVER CES DEFIS

### 15.8 Etre vigilant

Il faut que les programmes de préservation soient avertis des cadres juridiques dans lesquels ils opèrent, y compris leurs droits légaux, leurs contraintes et leurs obligations. Cette vigilance exige, le cas échéant, qu'on demande un avis juridique spécifique auprès d'une source compétente. La bonne intention de préserver des matériaux du patrimoine importants qui anime les programmes de préservation ne suffit pas : ils ont la responsabilité de mener à bien leur mission sans violer les droits légitimes des autres.

### 15.9 Une mobilisation

Les programmes de préservation doivent décider jusqu'à quel point ils se mobiliseront pour les questions de droits, en avançant des arguments en faveur d'une législation qui faciliteraient la préservation d'une vaste gamme de matériaux numériques.

Au minimum, les programmes de préservation doivent s'assurer que les parties intéressées sont averties des droits qu'exige une action de préservation efficace.

### 15.10 Trouver des solutions réalisables

Si trouver des solutions aux problèmes de droits risque d'être difficile, ces difficultés ne sont généralement pas insurmontables. Les résoudre exige le respect des intérêts légitimes des autres. Les solutions naissent souvent d'une approche coopérative qui reconnaît les besoins et les avantages mutuels. Les programmes de préservation y contribueront amplement en montrant que :

- une gestion saine des risques est possible ;
- il existe des moyens de répondre aux objectifs de préservation sans mettre en danger des intérêts commerciaux raisonnables ;
- grâce à leurs services de documentation et de métadonnées, les programmes de préservation peuvent promouvoir les connaissances de la communauté et encourager l'utilisation des produits des propriétaires des droits ;
- en sélectionnant les matériaux à préserver, les programmes de préservation confirmeront l'importance des enregistrements, des résultats de recherche et autres matériaux non publiés.

De nombreux programmes de préservation ont trouvé des manières satisfaisantes d'aborder les problèmes de droits, souvent en partenariat avec des propriétaires de droits. Ces modèles vont du simple accord avec des propriétaires de droits individuels (courants dans les archives de données et dans les archives sélectives de publications Web), à des partenariats à long terme entre de très grands éditeurs commerciaux et des bibliothèques nationales.

Ces modèles sont généralement fondés sur un mélange de droits transférés, gérés et conservés. Par exemple, le droit de stocker et de préserver le matériau peut être entièrement transféré, alors que le programme de préservation doit gérer étroitement l'accès et que le producteur conserve le copyright.

## PROBLEMES TECHNIQUES ET PRATIQUES

### 15.11 Les cadres juridiques

Il peut y avoir un certain nombre de cadres qui permettent aux programmes de préservation de s'approprier le droit de collecter et de préserver des matériaux numériques spécifiques. Parmi les plus courants figurent :

- un dépôt légal ou une législation de gestion des enregistrements ;
- des règles organisationnelles régissant une information d'entreprise ;
- des obligations contractuelles pour le dépôt des données ;
- des conditions en matière de subventions, d'emploi, ou un statut de membre d'une organisation ;
- des droits hérités d'une organisation par une autre ;
- des accords de licence négociés ou achetés. ;
- des droits consécutifs à la soumission volontaire d'un matériau à un programme de préservation.
- certains organismes de préservation saisissent et stockent des matériaux disponibles au public sur des sites Web en accès libre sans demander d'autorisation préalable. Certains agissent ainsi en revendiquant un "usage loyal" des matériaux du domaine public ; d'autres s'appuient sur une option d'"abandon" qui invite en général les propriétaires des droits à formuler une objection.

Il incombe au programme de préservation de déterminer, sur la base d'un avis juridique compétent, si l'une de ces méthodes ou d'autres est applicable, et ce qui est nécessaire en tant que protection juridique adéquate.

### 15.12 Quelques démarches courantes

Chaque situation réclame son propre jeu de dispositifs, mais les programmes de préservation sont tenus d'ajouter certaines mesures courantes comme :

- déterminer la situation juridique par rapport aux droits spécifiés par la législation, les règles organisationnelles existantes, ou les accords de licence ;
- identifier les droits qui seront nécessaires pour assumer pleinement des responsabilités de préservation ;
- identifier les propriétaires des droits afférents, et les autres parties prenantes qui ont une influence déterminante sur la nature des droits à négocier ;
- préparer une explication claire de ce qui est nécessaire et de la manière dont ce sera géré ;
- prendre contact avec les propriétaires des droits et négocier un régime de droits qui soit mutuellement acceptable ;

- enregistrer les responsabilités de gestion des droits dans des métadonnées qui soient clairement et sûrement adjointes aux matériaux concernés ;
- s'assurer que les responsabilités sont comprises par le personnel ;
- avoir des systèmes, des procédures et des outils sûrs en place pour contrôler l'accès et la copie, et pour vérifier la conformité ;
- si nécessaire, isoler les actions de préservation de tous les autres types d'accès et d'utilisation ;
- s'assurer que les utilisateurs comprennent leurs droits et leurs obligations juridiques.
- évaluer régulièrement les systèmes et les procédures pour vérifier qu'ils effectuent ce qu'ils sont supposés effectuer ;
- contrôler tous les facteurs qui peuvent déclencher une modification des droits comme l'écoulement d'une période de temps déterminée.

### **15.13 Négocier les conditions d'accès**

Le niveau d'accès que les programmes de préservation doivent chercher à atteindre dépendra de leur mission : il conviendra de soumettre certains matériaux du patrimoine numérique à un accès limité pour respecter la vie privée, la sécurité ou pour d'autres raisons, tandis qu'il semblera judicieux de rendre des matériaux publiés disponibles pour un accès continu grâce à un programme de préservation bien géré.

Certaines possibilités devraient avoir un effet favorable sur la négociation des conditions d'accès, entre autres :

- des restrictions géographiques d'accès, telle qu'une limitation aux utilisateurs sur site ;
- des limitations des possibilités de copie, comme l'utilisation d'un ordinateur autonome sans accès aux réseaux extérieurs ou des lecteurs de disquette ;
- une limitation du nombre d'utilisateurs qui puissent accéder au matériau à tout moment ;
- des seuils de temps permettant un accès non restreint après une période raisonnable d'exploitation commerciale ;
- une entente sur les facteurs déclenchant un transfert des droits d'accès, par exemple lorsque le matériau n'est plus disponible sur le site d'un éditeur ;
- un accès restreint pour les utilisateurs autorisés auxquels on demande de remplir des conditions particulières.

### **15.14 Gérer les droits**

Lorsque les droits ont été négociés, ils doivent être gérés en tant que responsabilité professionnelle essentielle du programme de préservation.

- Les programmes de préservation peuvent être appelés à traiter de vastes quantités de matériaux, aussi l'utilisation d'accords normalisés de licence englobant des classes de matériaux évitera-t-elle d'avoir à négocier et à gérer les droits élément par élément.
- Des outils de système pour gérer les droits sont disponibles et on attend qu'ils continuent d'évoluer. Ces outils enregistrent les conditions d'accès qui s'appliquent aux éléments individuels, enregistrent et filtrent les demandes d'utilisation, et rendent compte de l'utilisation. En choisissant des outils de gestion des droits, il importe de décider quels sont les outils appropriés pour soutenir une méthode équilibrée de gestion des droits.
- Il faut faciliter la mise en relation des utilisateurs avec les propriétaires de droits pour qu'ils négocient leurs propres autorisations, comme le droit de copie, là où il est de la responsabilité de l'utilisateur de le faire.
- Faciliter autant que possible l'accès autorisé peut dissuader un accès et une utilisation non autorisés.
- Encourager les créateurs à utiliser des logiciels au code source libre aidera à réduire les complications et les coûts dus à la négociation des droits avec les développeurs de logiciels brevetés.

### **POUR LES PROGRAMMES DE PRESERVATION DISPOSANT DE PEU DE RESSOURCES**

#### **15.15 Chercher des solutions efficaces**

Les problèmes de droits peuvent accroître considérablement les dépenses des programmes de préservation : tous les programmes ont donc intérêt à trouver des solutions efficaces et à éviter de s'exposer à des litiges. Les programmes disposant de peu de ressources devront particulièrement rechercher des contrats normalisés qui réduisent les coûts des accords de négociation de droits. Ils devront peut-être aussi accepter que la gestion des droits soit un facteur qui limite l'ampleur de leurs activités.

Dans cette perspective, peut-être tiendront-ils à limiter leurs activités à des matériaux qui présentent le moins possible de problèmes de droits, pour un certain nombre de raisons, entre autres :

- ils ont déjà l'autorisation ;
- les droits sont périmés (encore que cela soit improbable dans le cas des matériaux numériques pendant les prochaines décennies) ,
- la communauté du producteur a fortement intérêt à soutenir le programme de préservation ;
- ils disposent d'un avis juridique fiable qui soutient qu'un "usage loyal" ou d'autres dispositions constitueront une défense efficace ;

## ETUDES DE CAS

### 15.16 Etude de cas 1

Un centre d'archives de données qui travaille dans une discipline universitaire utilise une lettre d'accord normalisée à signer par les déposants. Celle-ci autorise le centre d'archives à faire une copie des données et à prendre toutes les mesures nécessaires de prévention, y compris procéder à des copies ultérieures dans tous les formats qu'il estimera les mieux adaptés pour fournir un accès fiable. Les déposants doivent indiquer s'il faut mettre des restrictions à l'accès, soit pour une période particulière de temps, soit pour des classes particulières d'utilisateurs, soit encore pour des types d'utilisation particuliers.

La période maximale pour un accès fermé est de dix ans. Le centre d'archives gère les droits manuellement, car les données ne sont pas disponibles en ligne : les demandes de l'utilisateur sont contrôlées par rapport aux enregistrements des métadonnées pour le matériel demandé avant que l'accès ne soit autorisé.

### 15.17 Etude de cas 2

Une bibliothèque d'Etat s'en remet à la législation du dépôt légal qui l'autorise spécifiquement à faire et à stocker des copies à des fins de préservation. Les conditions de copyright continuent de s'appliquer, de sorte que la bibliothèque informe les utilisateurs de la nécessité de demander l'autorisation au propriétaire du copyright avant de faire des copies. La bibliothèque négocie les restrictions d'accès avec les propriétaires des publications commerciales pour protéger leurs intérêts commerciaux pendant une certaine période de temps décidée en commun, généralement fixée à cinq ans, au cours de laquelle seule une utilisation exclusive *in situ* est autorisée. De nombreux propriétaires apprécient un accès moins restrictif parce qu'il élargit le public de leurs publications, tandis que d'autres exigent des périodes de restrictions plus longues. On utilise un système de métadonnées de gestion des droits pour enregistrer les restrictions et accepter ou rejeter automatiquement les demandes.

## REFERENCES - pour en savoir plus

### Références internes

*Travailler avec les producteurs*, voir aussi le chapitre 13.

*Les métadonnées*, voir aussi le chapitre 14.

### Références externes (tous les liens ont été consultés à la date de mars 2003)

De nombreuses archives de données utilisent des accords de licence avec les déposants pour donner une forme précise au transfert des droits. Voir, par exemple :

- Oxford Text Archive (2003), *License for depositors*. <http://ota.ahds.ac.uk/> , sous "OTA Publications".

Pour des exemples significatifs de négociations de gestion de droits donnant des résultats positifs, voir :

- IFLA and the International Publishers Association (juin 2002), Preserving the Memory of the World in Perpetuity: a Joint Statement on the Archiving and Preserving of Digital Information. <http://www.ifla.org/V/press/ifla-ipa02.htm>
- Koninklijke Bibliotheek (août 2002), National Library of the Netherlands and Elsevier Science Make Digital Preservation History. [http://www.kb.nl/kb/resources/frameset\\_kb.html?/kb/ict/dea/ltp/ltp-en.html](http://www.kb.nl/kb/resources/frameset_kb.html?/kb/ict/dea/ltp/ltp-en.html)
- Quelques ressources complémentaires :
- CEDARS Project (2002), CEDARS Guide to Intellectual Property Rights. <http://www.leeds.ac.uk/cedars/guideto/ipr/guidetoipr.pdf>
- Kavcic-Colic, Alenka (2002), Archiving the Web: Some legal Aspects, 68th IFLA Council and General Conference, Glasgow. <http://www.ifla.org/IV/ifla68/papers/116-163e.pdf>

## Chapitre 16. Protéger les données

### INTRODUCTION

#### 16.1 Les objectifs

Ce chapitre permettra au responsable de programme de comprendre combien il est important de maintenir un contrôle strict sur l'intégrité des données sous-tendant les objets numériques. Ceux qui sont engagés dans l'exécution pourront utiliser des informations contenues dans ce chapitre qui leur serviront de base pour discuter des besoins nécessaires avec les spécialistes des techniques de l'information ou des prestataires de services.

#### 16.2 En bref

La protection des données est la préoccupation fondamentale de tous les programmes de préservation. Pour de nombreux programmes, l'authenticité est également un point essentiel, qui dépend de l'intégrité permanente des données, et de leur identification lisible et continue. Les stratégies de protection de données incluent une attribution de la responsabilité, une infrastructure technique, un entretien, un transfert des données, un stockage adéquat des supports de données, une sauvegarde, une sécurité du système et un plan anti-catastrophe. L'authenticité dépend aussi d'une documentation lisible des origines et de l'histoire des matériaux numériques.

### LES PRINCIPAUX PROBLEMES DE GESTION

#### 16.3 Le stockage et la protection des données

Les données doivent être stockées. S'il convient de concentrer les soins de la préservation sur la meilleure façon de représenter les objets numériques tels qu'ils ont été originalement prévus, il ne faut pas oublier que les objets numériques ont une forme sous-jacente en tant que données. C'est en tant que données qu'elles doivent être stockées, gérées et protégées si l'on veut qu'un objet numérique soit disponible pour être présenté à un utilisateur.

#### 16.4 L'authenticité

Les matériaux du patrimoine sont souvent estimés, au moins en partie, en fonction de leur authenticité : jusqu'à quel point on peut être sûr qu'ils correspondent vraiment à ce qu'ils sont censés être pour les enregistrements d'archivage, les données scientifiques, et bien d'autres sortes de matériaux numériques, la confiance en leur authenticité permanente est d'une importance capitale, car sans elle ils sont quasiment sans valeur.

L'authenticité vient de ce qu'on peut se fier à la fois à l'*identité* d'un objet - il est ce qu'il dit être, et n'a été confondu avec aucun autre - et à son *intégrité* - il n'a subi aucune modification qui en ait modifié le sens.

Le maintien tout à la fois de l'identité et de l'intégrité implique des liens prolongés et documentés entre un objet tel qu'il a été créé originellement et tel qu'il est présenté actuellement.

Evaluer, conserver et fournir la preuve d'une authenticité permanente sont des responsabilités essentielles pour la plupart des programmes de préservation.

## **16.5 Les menaces qui pèsent sur l'authenticité**

L'authenticité peut être compromise par :

- des menaces sur l'identité. La perte de certitude quant à la manière dont un objet se distingue d'autres objets porte atteinte à l'authenticité. Cela peut résulter d'une confusion dans l'identification des données, de modifications apportées aux identifiants, ou d'une carence de la documentation des relations entre les différentes versions ou copies ;
- des menaces sur l'intégrité. Des modifications du contenu de l'objet lui-même peuvent aussi porter atteinte à l'authenticité. La plupart de ces modifications proviennent de menaces qui pèsent sur l'objet au niveau des données.

La nature des matériaux numériques, et la manière dont ils sont gérés pour la préservation et l'accès posent tous deux des défis :

- les matériaux numériques peuvent être facilement modifiés, avec ou sans intention frauduleuse, et même sans la moindre intention ;
- la modification apportée peut ne pas être évidente ;
- les processus de préservation signifient presque toujours l'apport de modifications : le transfert de données d'un système à un autre, d'un support à un autre, l'ajout ou la mise à jour de métadonnées, la création de nouvelles copies qui nécessitent de nouveaux noms de fichier, la modification des moyens de présentation comme les changements de technologies, etc.

## **16.6 Les menaces qui pèsent sur l'intégrité des données**

Parmi les menaces qui pèsent couramment sur l'intégrité permanente des données et auxquelles les programmes de préservation risquent d'être en butte figurent :

- une génération "naturelle" d'erreurs qui surgissent dans les systèmes de stockage numérique.
- la défaillance des supports. La plupart de ceux-ci ont une vie exploitable assez courte avant de se détériorer au point de ne plus être fiables pour le stockage des données.
- une attaque malveillante qui peut provenir de pirates du système, de virus, du personnel ou d'intrus extérieurs communiquant avec le système de stockage.
- les dommages collatéraux provoqués par des actes malveillants comme les attaques terroristes, les faits de guerre ou les manifestations civiles qui s'en prennent aux bâtiments ou aux approvisionnements en électricité.
- les actes commis par inadvertance par des membres du personnel ou des visiteurs comme l'arrêt de l'électricité, la mise à la poubelle de disquettes ou de bandes, ou le reformatage des dispositifs de stockage.
- les catastrophes "naturelles", comme les incendies, les inondations ou l'effondrement du bâtiment.
- la faillite commerciale.

La probabilité et les conséquences de ces risques et d'autres encore varieront d'une situation à l'autre. Toutefois, on peut partir du principe qu'il faut tenir compte de tous ces risques.

## DES PRINCIPES POUR RELEVER CES DEFIS

### 16.7 Quel est le degré d'authenticité nécessaire ?

Lorsque les matériaux numériques ont une valeur en tant qu'enregistrements qui apportent un témoignage sous une forme ou sous une autre, l'authenticité est de toute première importance. Tous les matériaux numériques ne sont pas créés ou sélectionnés pour fournir un témoignage : par exemple ils peuvent refléter une expression créative, un débat d'idées ou le désir de divertir ou d'être diverti. Même pour ces matériaux, la question de l'authenticité se pose dans la mesure où l'on doit protéger l'intégrité de l'oeuvre ou des idées de leurs créateurs.

Finalement, les programmes de préservation doivent décider de la somme qu'ils investiront pour garantir la fiabilité de l'authenticité du matériau qui leur est confié, en gardant à l'esprit que l'identité de l'objet et l'intégrité des données sont des responsabilités fondamentales.

### 16.8 Le rôle de la protection des données

La protection des données doit jouer un rôle fondamental dans tous les programmes de préservation, pour deux raisons :

- afin qu'il y ait un objet numérique auquel l'utilisateur puisse accéder. C'est une nécessité primordiale : si les données sont perdues ou gravement altérées, il peut être totalement impossible de représenter l'objet numérique prévu et l'on devra considérer alors que le processus de préservation a échoué ;
- afin que l'intégrité des données puisse être maintenue sans truquage ni altération de sorte que les utilisateurs puissent se fier à l'authenticité de l'objet re-présenté.

### 16.9 Le rôle de la documentation

La documentation joue également un rôle essentiel pour deux raisons :

- en expliquant les liens entre les objets et en faisant clairement la distinction entre eux, elle fournit une preuve de l'identité ;
- en montrant quelles modifications, s'il y en a, sont survenues, avec quelle autorisation et dans quel dessein, elle fournit un fichier de vérification pour certifier l'authenticité.

### 16.10 Les responsabilités pour maintenir l'authenticité

Il ne serait pas réaliste d'espérer une garantie d'authenticité totalement objective - il peut toujours y avoir un élément de confiance ou de jugement subjectif lorsque l'on décide que l'authenticité a été suffisamment prouvée -, toutefois il semble raisonnable d'attendre que les programmes de préservation numériques assument trois responsabilités :

- ils doivent estimer si l'authenticité démontrée est d'une importance vitale pour la valeur durable du matériau ;

- ils doivent protéger le matériau qui leur est confié des modifications qui altéreraient son sens. (Ce qui permet des modifications externes comme les nouvelles interprétations, sans permettre des modifications internes qui altéreraient son sens) ;
- ils doivent documenter les relations sur lesquelles repose le niveau requis d'authenticité. Celles-ci incluent les relations entre les objets et leurs identifiants ; entre les objets et leur producteur ; entre les différents objets ; entre l'objet et la manière dont il a été géré.

### **16.11 Les stratégies de protection des données**

D'autres sortes de matériaux numériques peuvent survivre à des périodes de négligence, mais pas les données numériques. Les objets numériques exigent des stratégies bien planifiées, bien gérées et prolongées pour protéger les données comme bases minimales de continuité. Les stratégies nécessaires comprennent généralement :

- une attribution claire des responsabilités ;
- l'apport d'une infrastructure technique appropriée, y compris les systèmes, les dispositifs de stockage, et les supports pour accomplir le travail ;
- des programmes d'entretien, de soutien et de remplacement du bien pour les systèmes ;
- un transfert régulier des données sur des supports nouveaux pour s'assurer qu'elles ne sont pas menacées par une détérioration des supports ou des modifications dans le matériel d'accès ;
- un stockage et des conditions de manipulation appropriés pour les supports ;
- un niveau élevé de redondance pour se prémunir contre la défaillance de toute copie ou composant unique ; y compris des régimes de sauvegarde appropriés ;
- un niveau élevé de sécurité du système, y compris des contrôles sur l'accès aux données stockées ;
- un plan anticatastrophe.

Ces stratégies sont plus détaillées ci-dessous dans les problèmes techniques et pratiques.

## **PROBLEMES TECHNIQUES ET PRATIQUES**

### **16.12 Faire appel aux prestataires de services**

De toutes les responsabilités des programmes de préservation, le stockage et la protection des données sont sans doute celles pour lesquelles il est le plus facile de trouver des prestataires de service tiers qui soient appropriés. Compte tenu des investissements importants qu'exigent l'équipement et les compétences, cela peut être une solution de rechange attrayante pour la gestion interne des données. Mais quand on sait le caractère vital de la protection des données, il est indispensable que les programmes de préservation s'assurent que tous les services sous contrat soient en mesure de remplir les tâches de surveillance et de contrôle nécessaires.

### **16.13 Les aspects pratiques des stratégies de protection des données**

On utilise un assortiment de stratégies assez normalisées pour gérer des données en vue d'un stockage à long terme. La plupart se fondent sur l'hypothèse qu'il faut préserver seulement les données et pas le support de données lui-même.

- L'attribution de la responsabilité. On doit assigner à quelqu'un une responsabilité sans équivoque pour gérer le stockage et la protection des données. C'est une responsabilité technique qui réclame une gamme particulière de compétences et de connaissances ainsi qu'une expérience de la gestion. Excepté pour de très petites collections, le stockage et la protection des données exigent qu'on y consacre des ressources, que l'on élabore un plan approprié et que l'on puisse rendre des comptes en ce qui concerne ces stratégies.
- Une infrastructure appropriée pour faire le travail. Les données doivent être stockées et gérées avec des systèmes adéquats et sur un support adéquat. Il y a des systèmes de gestion du bien numérique ou des systèmes de stockage d'objets numériques disponibles qui répondent aux exigences des programmes de préservation numérique. Une fois les exigences déterminées, on doit en discuter à fond avec les fournisseurs éventuels. Des systèmes et des supports différents sont adaptés à des besoins différents et ceux qui sont choisis pour les programmes de préservation doivent correspondre à leurs objectifs. Le système global doit avoir des capacités appropriées, notamment :
  - une capacité de stockage suffisante. On peut l'augmenter avec le temps, mais le système doit être à même de gérer la quantité de données prévue pour être stockée pendant leur cycle de vie ;
  - en tant que capacité fondamentale, le système doit pouvoir, si nécessaire, dupliquer les données sans perte, et transférer les données sur des supports nouveaux ou "rafraîchis" sans perte ;
  - une fiabilité démontrée et un soutien technique pour traiter rapidement les problèmes ;
  - la capacité d'affecter un deuxième nom de fichier dans un système de désignation de fichier adapté à son architecture de stockage. Les systèmes sont fondés autour d'objets nommés. Cela peut imposer des contraintes sur la manière dont les objets sont nommés à l'intérieur du stockage ; par exemple, les systèmes sur disquette peuvent imposer une structure hiérarchique de répertoire à des noms de fichier existants, différents de ceux qui seront utilisés sur un système sur bande. Le système doit permettre, ou de préférence effectuer, l'affectation d'un deuxième nom de fichier et d'identifiants existants imposés par le système ;
  - la capacité de gérer un stockage redondant ;
  - la recherche d'erreurs. Un niveau de recherche d'erreurs automatisée est une chose normale dans la plupart des stockages informatiques. Comme il faut conserver les matériaux du patrimoine pendant de longues périodes de temps, souvent avec une utilisation humaine très faible, le système doit pouvoir détecter les modifications et les pertes de données et prendre les mesures appropriées.
- L'infrastructure technique doit également inclure des moyens de stocker les données et de relier d'une manière fiable les métadonnées aux objets numériques stockés. Les opérations importantes découvrent souvent qu'elles ont besoin d'établir des systèmes de gestion d'objet numérique qui sont reliés, mais séparés de leur système numérique de mémoire de masse, afin de faire face à la

gamme de processus mis en jeu, et de permettre aux métadonnées et aux interfaces de travail d'être modifiées sans avoir à modifier la mémoire de masse.

- Les options larges actuellement disponibles pour des supports de stockage à une grande échelle sont détaillées ci-dessous dans le tableau 16.1 :

---

## Support

---

Disque magnétique (par exemple disque dur)

---

Bande magnétique

---

Disque optique (CD, DVD)

---

## Accès aux données

---

Accès aléatoire rapide

---

Stockage linéaire, prend donc beaucoup plus de temps pour chercher les données et y accéder aux données

---

Accès aléatoire rapide, mais plus lent que le disque magnétique

---

## Permet une modification des données ?

### Oui

---

Généralement non - la "réinscription" exige que les données soient écrasées

---

Oui, sur certains produits

---

## Capacités actuelles de stockage par unité

---

Jusqu'à 200 gigaoctets

---

Jusqu'à 200 gigaoctets

---

Jusqu'à 4 gigaoctets

---

**Vitesse d'augmentation de la capacité**

---

Double tous les 12 à 18 mois

---

Double tous les 12 à 18 mois

---

Lente car ce support n'est pas utilisé pour les archives ou les sauvegardes très importantes

---

**Durée d'utilisation attendue pour une seule unité**

---

Environ cinq ans

---

Environ cinq ans

---

Large fourchette d'environ cinq ans pour les produits bas de gamme à plusieurs décennies pour les produits haut de gamme

---

**Autres commentaires**

---

Supports généralement non amovibles

---

Supports portatifs adaptés à la sauvegarde

---

Supports portatifs

---

Coût de l'unité faible ; matériel grand public de faible coût largement disponible

---

**Tableau 16.1** - Comparaison entre les supports de données à grande échelle

- Les programmes d'entretien, de soutien et de remplacement. Les composants du système doivent généralement être remplacés au bout de quelques années. Le matériel, en particulier, a une durée de fonctionnement d'environ cinq ans avant qu'il ne devienne difficile de se procurer une assistance technique. Les supports de stockage ont également besoin d'un "rafraîchissement" (réécriture des données) et d'un remplacement périodique par de nouveaux supports.

La nécessité de remplacer les systèmes de stockage entraîne des frais courants importants, qui couvrent l'équipement lui-même ainsi que les processus d'approvisionnement et de transferts de données qui précèdent et suivent l'installation d'un nouvel équipement. Ces coûts peuvent être intégrés dans un plan de budget à long terme.

Il faut prendre en compte le coût de remplacement des supports de données mais savoir aussi que les médias de remplacement offrent, de leur côté, une capacité de stockage accrue. Malheureusement, les économies réalisées sont généralement contrebalancées par l'accroissement de la quantité de données à stocker.

Le marché des systèmes de stockage et de gestion des données s'étend bien au-delà des programmes de préservation, aussi y a-t-il de bons produits COTS (Commercial Off the Shelf) disponibles sur le marché. L'utilisation de la technologie COTS est certainement la méthode la plus facilement gérable, la plus compétitive et offrant le moins de risques, car l'assistance technique et les mises à jour sont fournies par les fabricants sur un marché concurrentiel. Les normes sont d'un usage assez répandu sur le marché du stockage pour autoriser un mélange et un assortiment de produits venant de divers fabricants, de sorte que l'on dispose de nombreuses voies de mise à jour et de remplacement lorsque c'est nécessaire.

- Un transfert régulier des données sur de nouveaux supports. Les systèmes de stockage reposent sur une reproduction sûre et complète des données, plutôt que sur des supports durables, pour la protection des données. Il faut copier celles-ci d'un support à un autre pour éviter qu'elles ne subissent les conséquences d'une détérioration du support. Au fur et à mesure que de nouvelles formes de support démontrent leur utilité dans les systèmes de stockage, les données seront transférées sur ceux-ci à partir des supports de type plus ancien. Cette opération doit être effectuée avant que le matériel ou le logiciel nécessaires pour consulter les données ne soit abandonné.

La planification des transferts de données est un défi sur le plan de la gestion, quel que soit le système utilisé. Par exemple, un petit centre d'archives peu utilisé qui stocke des données dans des CD rangés sur des étagères doit surveiller le vieillissement et l'état des CD ainsi que les signes indiquant qu'il faudra remplacer la technologie du CD par une autre. Des systèmes de mémoire de masse plus perfectionnés automatisent généralement les décisions relatives au transfert régulier des données entre supports, mais les responsables auront toujours à décider du moment où ils devront remplacer les supports par de nouveaux médias, et où la technologie sous-jacente se révèle dépassée.

- Un stockage et des conditions de manipulation appropriés pour les supports. Les supports de données numériques doivent être stockés dans des conditions qui n'accélèrent pas leur vitesse de détérioration.

Les principaux risques qui guettent les supports de données sont une température et une humidité excessives : ils mettent le support en danger ; la poussière ou d'autres particules peuvent boucher l'accès aux données ; et dans le cas de matériaux codés optiquement, la lumière peut endommager les données inscrites optiquement. Les bandes de données modernes ont une telle force coercitive, qu'un effacement accidentel par un champ magnétique n'est pas un grand risque.

On peut intégrer les bandes de données magnétiques dans un système de stockage numérique. Celui-ci, comme il se doit, sera logé dans un local pour ordinateur hors poussière avec une surveillance de la température et du taux d'humidité relative, fixés respectivement à 18°C et à 40 %, un flux régulier d'air propre et sans poussière, et un nettoyage quotidien pour éviter la contamination. Les conditions de température ne devront pas varier de plus de 2° et de celles du taux d'humidité relative de plus de 10 % pour une période donnée de 24 heures.

Les bandes de données magnétiques emmagasinées en vue d'une durée de vie optimale du support (hors de l'environnement du local de l'ordinateur) devront être stockées dans des conditions moins rigoureuses, à une température comprise entre 10°C et 18°C avec une tolérance journalière ne dépassant pas 1°C, et un taux d'humidité relative situé entre 30 et 40 % avec une tolérance ne dépassant pas 3 %.

Les supports optiques, comme les CD-ROM enregistrables, seront stockés dans des conditions similaires, mais dans un environnement sombre en raison de leur sensibilité à la lumière.

Certains prétendent que des températures très basses (voisines ou inférieures à 0°C) peuvent nuire à l'espérance de vie de certains supports, toutefois cela n'est pas prouvé.

- Les régimes de redondance et de sauvegardes. On ne saurait trop souligner l'importance des régimes de redondance et de sauvegarde : ils sont essentiels pour tous les programmes de préservation numérique en tant que garantie fondamentale contre les détériorations ou la perte de toute copie unique.

Certes, le stockage de copies multiples des mêmes données offre une protection contre les défaillances, mais les programmes de préservation doivent aussi envisager les risques de voir se produire une catastrophe qui porte atteinte à toutes les copies stockées dans un même site. Stocker des copies dans des sites différents est une nécessité fondamentale ; pour éviter les conséquences de catastrophes régionales comme les inondations, les tremblements de terre, les incendies et la guerre, les programmes doivent prendre en compte la nécessité de stocker des copies de sauvegarde supplémentaires des données importantes à l'extérieur de leur propre région.

Les programmes de préservation peuvent aussi avoir besoin d'adapter des calendriers de sauvegarde normale pour que les données de préservation à conserver soient rafraîchies (c'est-à-dire réécrites) et non pas refrappées avec de nouvelles données.

- La sécurité du système. Des contrôles de sécurité sont indispensables pour s'assurer que les données stockées ne sont exposées qu'à des processus contrôlés et autorisés. Les mesures de sécurité aux normes TI pour les biens de l'information d'importance capitale sont pleinement applicables et rigoureusement nécessaires.
- Le plan anticatastrophe. Il faut mettre en place et tester régulièrement des plans de récupération de catastrophe aux normes TI. Ils peuvent inclure des dispositifs réalistes pour essayer de récupérer des données sur des supports endommagés, mais la récupération des données est coûteuse et peu sûre ; on doit la considérer comme une solution de remplacement fort peu satisfaisante à des dispositifs de "récupération à partir de la sauvegarde" voulus.

## **16.12 Gérer les risques**

Le tableau 16.2 présente une analyse simplifiée des risques liés à certaines des menaces les plus courantes qui pèsent sur les données en stockage.

---

## La menace

---

Génération "naturelle" d'erreurs

---

Défaillance de support

---

Attaque malveillante : pirates, virus, intrus

---

Domage collatéral provenant d'autres attaques non dirigées contre le système

---

Des actes faits par inadvertance, par exemple couper le courant, jeter les fichiers, reformater les dispositifs de stockage

---

Catastrophes naturelles, par exemple les incendies, les inondations

---

Faillite commerciale

---

## Ce qui est atteint

---

L'intégrité des données

---

L'intégrité des données, l'identité du fichier

---

L'intégrité des données, l'identité du fichier, les biens d'équipement

---

L'intégrité des données, l'identité du fichier

---

L'intégrité des données, l'identité du fichier, les biens d'équipement

---

L'accès aux données

---

**Probabilité**

---

Presque certaine

---

Certaine pour la plupart des supports

---

Presque certaine pour les archives mises en réseau

---

Variable selon la situation

---

Probable, sauf s'ils sont gérés

---

Très probable à long terme

---

Variable

**Vitesse de l'attaque**

---

Progressive

---

Progressive

---

Probablement soudaine

---

Probablement soudaine et inattendue

---

Probablement inattendue

---

Probablement soudaine mais il peut y avoir une période d'avertissement

---

Peut être progressive ou soudaine

---

## **Conséquences**

---

Les données peuvent ne pas fonctionner ; peut empêcher la récupération des données graves ; les données peuvent être illisibles et non récupérables

---

Probablement graves ; peuvent inclure une réécriture ou l'altération des données

---

Probablement graves, et dépassant les mesures normales de sécurité

---

Varié du dommage à la catastrophe

---

Peuvent être localisées et perte minimale ou totale

---

Perte d'accès probable lorsque les supports sont vidés ou réutilisés

---

## **Options de prévention**

---

Recherche d'erreur, correction d'erreur, rafraîchissement des données et transfert

---

Utilisation de produits haut de gamme ; utilisation de supports plus stables ; vérifier fréquemment l'état ; transférer les données avant la fin de durée de vie du support

---

Mesures de sécurité, logiques et physiques ; barrières de sécurité ; contrôles d'accès, mettre les données hors ligne

---

Données de sauvegarde ; assurer l'accès aux sauvegardes

---

Données de sauvegarde ; formation du personnel et contrôle d'accès physique

---

Plan d'alerte ; zones de stockage bien situées ; sauvegarde hors site

---

Planification et gestion commerciales ; dispositifs de continuité/succession ; identification claire des biens importants.

---

*Tableau 16.2 - Exemple d'analyse de risque des menaces contre la protection des données*

## POUR LES PROGRAMMES DE PRESERVATION DISPOSANT DE PEU DE RESSOURCES

### **16.13 Les responsabilités essentielles**

La protection des données est une responsabilité si importante que même les programmes disposant de peu de ressources doivent lui donner un rôle prioritaire. L'analyse de risque simplifiée qui figure ci-dessus laisse entendre qu'il existe des zones de risque moins élevé pour certains programmes. Elle suggère également que l'on peut réduire certains risques à condition de réduire le degré de rapidité de l'accès. Cette solution peut convenir pour certaines collections.

### **16.14 Donner des priorités**

On peut privilégier certaines parties de la collection en leur assurant une protection supplémentaire, et offrir à des données moins importantes une protection moindre (comme des sauvegardes moins fréquentes, l'utilisation de supports de moins bonne qualité, un transfert de données sur de nouveaux supports moins fréquent).

## ETUDES DE CAS

### **16.15 Etude de cas 1**

Un dispositif particulier qui emploie la redondance possède des données sur bande en trois copies : la première conservée près du réseau dans une bibliothèque rattachée au système, la deuxième autonome mais sur site et la troisième hors site. Toutes les copies sont identiques et le système les conserve telles quelles. Pour l'accès, on fait une copie provisoire sur des disques organisés comme un RAID (Redundant Array of Inexpensive Disks) ("batterie en double emploi de disques bon marché") dans lequel la défaillance de l'un des disques est compensée par des copies sur les autres disques. Pour effectuer la redondance du support, il peut aussi y avoir une copie stockée hors site sur support optique.

### **16.16 Cas d'étude 2**

Un centre d'archives d'enregistrement documentant les transactions commerciales du gouvernement s'efforce de certifier l'authenticité de chaque enregistrement qu'il stocke. Tous les enregistrements scannés d'après les originaux non numériques comportent une déclaration attestant qu'il s'agit de vraies copies ; les enregistrements numériques saisis à partir des systèmes de conservation des enregistrements électroniques comprennent des contrôles de vérification créés par le système. Tous les processus qui pourraient occasionner des modifications involontaires ou non autorisées sont documentés dans les métadonnées de préservation reliées à l'enregistrement.

Une bibliothèque régionale qui rassemble des publications numériques utilise des vérifications du contrôle de la qualité pour garantir que les fichiers qu'elle saisit coïncident avec l'exemplaire restant sur le site de l'éditeur. Elle documente les processus qu'elle applique au matériel, et contrôle tous les risques importants qui pourraient menacer l'intégrité des données, mais elle accepte que certains processus conduisent à des éléments qui diffèrent de leur apparence originale quand ils seront représentés à l'avenir. Elle ne peut pas certifier que les copies qu'elle

présente sont authentiques, mais prétend que ses processus fournissent des bases raisonnables pour les accepter en tant que copies archivées et gérées à des fins de recherche.

## REFERENCES - pour en savoir plus

### Références internes

*Relationship between data and presented digital objects*, voir également Comprendre la préservation numérique : chapitre 7.

### Références externes (tous les liens ont été consultés à la date de mars 2003)

On trouvera quelques points de vue intéressants dans :

- Gladney Henry M., *Digital Document Quarterly*. <http://home.pacbell.net/hgladney/ddq.htm>
- Graham Peter S. (2000), *Authenticity in a Digital Environment*, Council on Library and Information Resources. <http://www.clir.org/pubs/reports/graham/intpress.html>
- InterPARES Project (2002). *The long-term Preservation of Authentic Electronic Records: Finding of the InterPARES Project*. <http://www.interpares.org/book/index.htm>

On trouvera des informations techniques sur les dispositifs de stockage des données dans :

- Bogart John Van (1995). *Magnetic Tape Storage and Handling: A Guide for Libraries and Archives*, Council on Library and Information Resources, Washington, DC. <http://www.clir.org/pubs/reports/pub54/index.html>
- CoOL [*Conservation OnLine*]: *electronic storage media*. <http://palimpsest.stanford.edu/bytopic/electronic-records/electronic-storage-media/>
- Kodak Professional (sans date), *Permanence and Handling of CDs*. <http://kodak.com/global/en/professional/products/storage/pcd/techInfo/permanence.jhtml>
- Library of Congress (éd. rev. 2002). *Cylinder, Disc and Tape Care in a Nutshell*. <http://www.loc.gov/preserv/care/record.html>

## Chapitre 17. Maintenir l'accessibilité

### INTRODUCTION

#### 17.1 Les objectifs

Ce chapitre vise à expliquer le contexte du maintien de l'accès, et ce qui est nécessaire pour le soutenir, tout en fournissant une base pour comparer une gamme de stratégies couramment proposées.

#### 17.2 En bref

Les modifications des logiciels et du matériel aboutissent à une perte des moyens d'accès des matériaux du patrimoine numérique. C'est généralement le défi majeur que rencontrent la plupart des programmes de préservation. En utilisant les explications de la relation entre les objets numériques et leurs moyens d'accès et en prenant en compte ce qui doit être présenté à l'utilisateur lorsqu'on fournit l'accès, les responsables de programme doivent se décider sur des stratégies qui garantiront l'accès chaque fois qu'on en aura besoin. Les stratégies, qui sont susceptibles de varier avec le temps et en fonction des besoins, continuent d'évoluer. Les stratégies étudiées ici sont regroupées en trois catégories : celles qui sont fondées sur un investissement de ressources effectué très tôt dans le cycle de vie des matériaux numériques, celles qui ont une efficacité à court, à moyen et jusqu'à long terme, et certaines stratégies de remplacement "non numériques" et de "non-préservation".

### LES PRINCIPAUX PROBLEMES DE GESTION

#### 17.3 Pourquoi des passerelles d'accessibilité sont nécessaires

Préserver la possibilité d'accéder au matériau numérique est l'objectif principal des programmes de préservation. Fondés sur des données et des métadonnées préservées, et en utilisant les outils d'accès des logiciels et du matériel, les objets numériques doivent être représentés aux utilisateurs sous une forme compréhensible. Cela devra être fait à n'importe quel moment dans l'avenir lorsqu'on en a besoin, en utilisant les technologies d'accès disponibles dans ce futur.

Comme les objets numériques reposent sur des formes combinées de technologie pour la présentation, la capacité de les représenter à une date ultérieure est brisée ou perdue au fur et à mesure que les technologies se modifient. Ce phénomène de modification des technologies d'accès est si courant que c'est presque une marque distinctive des matériaux numériques stockés.

#### 17.4 Des cadres temporels pour la préservation

En raison de la vitesse des changements technologiques, l'horizon de la perte se rapproche pour de nombreux matériaux numériques actuellement disponibles. Il est difficile, sinon impossible, de rendre accessibles avec les technologies d'aujourd'hui certains matériaux créés avec des technologies qui étaient courantes il y a moins de dix ans.

L'ultime objectif consiste à trouver les moyens de garantir l'accès à n'importe quel moment de l'avenir sur le long terme, il n'en reste pas moins nécessaire d'assurer l'accessibilité sur le court terme.

## **17.5 Définir les niveaux de perte acceptables**

Les programmes de préservation risquent fort d'être confrontés à des scénarios qui exigent de déterminer les niveaux de pertes acceptables et inacceptables.

Une fidélité absolue à la présentation originale des matériaux numériques sera difficile dans n'importe quel cas ; nombre de stratégies passées actuellement en revue peuvent impliquer des pertes, y compris une éventuelle perte du contenu, la perte de la présentation d'origine, ou la perte de quelques fonctions originales.

Ces pertes peuvent être une conséquence involontaire de la stratégie choisie (elle est courante lorsque l'on fait passer des fichiers dans un nouveau format), ou le résultat attendu du choix adopté pour réduire les coûts de préservation (comme lorsque l'on jette des liens ou des éléments dynamiques de pages Web). Elles peuvent même être inhérentes aux objectifs de préservation du programme (comme le retrait des fonctions de modification des documents sauvegardés en tant qu'enregistrements statiques).

Dans ces scénarios et les scénarios similaires, le programme réclame certains moyens pour déterminer quelles pertes seront acceptables.

### DES PRINCIPES POUR REpondre A CES DEFIS

## **17.6 La responsabilité des programmes de préservation**

Les programmes de préservation doivent trouver des moyens pour contourner la menace que représentent les changements et l'obsolescence des technologies, s'ils veulent atteindre leur objectif principal, qui est de maintenir la continuité de l'accès.

## **17.7 Reconnaître les éléments qu'il faut préserver**

De nombreuses collections contiennent des versions multiples des mêmes matériaux, comme des images numériques de haute qualité et leur version de qualité inférieure, des versions dérivées fournies pour un accès facile au réseau. Les programmes de préservation doivent déterminer la version, ou des versions, qu'il faut conserver, et celles qui pourront être créées à nouveau ultérieurement.

## **17.8 Reconnaître les éléments qu'il faut conserver**

Afin de définir des niveaux de perte acceptables et inacceptables, les programmes de préservation doivent définir les éléments essentiels qu'il faut conserver. Comme il l'a été dit précédemment (au chapitre 12), les programmes ont besoin de cette information pour :

- choisir la stratégie la plus apte à conserver ces éléments ;
- choisir la stratégie la plus compétitive ;
- évaluer si leur stratégie a réussi.

Fixer des objectifs de préservation à ce niveau exige une étude approfondie du matériau pour comprendre pourquoi il existe, comment il fonctionne, et ce qu'un utilisateur pourrait attendre et faire d'une copie préservée.

Une fois les éléments essentiels définis, la tâche du programme de préservation consiste à trouver, et à ne pas cesser de trouver, des combinaisons de données, de logiciels et de matériel qui représenteront ces éléments aussi précisément qu'il convient.

### 17.9 La relation entre les données et le logiciel

Il y a toujours une relation de dépendance entre les données et le logiciel : toutes les données réclament un logiciel afin d'être présentées à un utilisateur sous une forme compréhensible. Le degré de dépendance a des conséquences importantes :

- certains objets sont relativement indépendants d'un logiciel spécifique ; par exemple des ensembles de données brutes, un texte en clair ou en mode point comme l'ASCII peut être présenté en utilisant une gamme d'outils logiciels tout à fait élémentaire ;
- certains objets dépendent de logiciels plus complexes mais d'un usage général ou largement disponibles ; par exemple l'HTML, des formats d'image normalisés comme le TIFF et d'autres formats conçus pour fonctionner sur des plate-formes interchangeables ;
- certains objets dépendent d'un logiciel d'application spécifique et ne sont pas conçus pour fonctionner en dehors de leur environnement d'exploitation original - même si les fabricants fournissent souvent des outils qui permettent de les lire ou de les utiliser d'une façon plus restreinte ; par exemple des formats de traitement de texte, des feuilles de calcul, certaines bases de données, du dessin et des formats de cartographie GIS ;
- certains objets sont par essence logiciels ; comme par exemple les fichiers exécutables, les programmes logiciels ;
- de nombreux matériaux complexes contiennent des combinaisons d'objets avec des niveaux différents de dépendance logicielle.

Le degré de dépendance logicielle peut limiter le choix des stratégies disponibles. Par exemple, les objets "données" ou "type document" peuvent être présentés efficacement par une gamme de logiciels, tandis que les objets "logiciels" offrent bien moins d'options pour conserver l'accès une fois que leur environnement d'exploitation original a été perdu.

### 17.10 Choisir les stratégies pertinentes

Il n'y a, jusqu'à présent, aucune solution pratique universellement applicable au problème de l'obsolescence technologique pour les matériaux numériques. On a proposé plusieurs méthodes mais il est peu probable que se dégagera une solution unique offrant des moyens d'accès compétitifs pour tous les matériaux, à toutes fins et à tout moment. A ce stade, il est raisonnable pour les programmes de préservation de rechercher plusieurs stratégies, surtout s'ils sont responsables d'une gamme de matériaux sur des périodes prolongées.

Il est important de prendre dès maintenant des mesures efficaces qui, même si elles sont limitées, préserveront l'accès pour le "futur gérable", tout en envisageant aussi quelles méthodes à long terme semblent les plus pratiques.

Les stratégies à long terme actuellement privilégiées semblent être les suivantes : l'utilisation de *normes* pour le codage, l'organisation et la description des données qui sont supposées rester identifiables pendant de longues périodes ; l'*émulation* d'un logiciel ou d'un matériel dépassé dans un nouvel environnement ; et le *passage* de données d'une technologie d'exploitation à une autre.

Toutes ces stratégies ont montré qu'elles fonctionnaient dans certaines circonstances sur des périodes de temps limitées. Forcément, elles n'ont pas fait leurs preuves face à des menaces inconnues sur des siècles de changement. Mais elles sont actuellement appliquées dans la gestion des données, et il semble probable que, conjuguées entre elles, elles continuent d'être recherchées et proposées pour une préservation à grande échelle sur le long terme.

### **17.11 Les principes qui sous-tendent les méthodes actuelles**

En cherchant les moyens de maîtriser les changements technologiques, la plupart des méthodes proposées jusqu'alors se fondent sur un des principes suivants ou sur plusieurs :

- la conversion des données sous une forme humainement lisible sur un support facile à conserver (comme le papier, le film ou des supports métalliques stables) ;
- la création ou la conversion de données dans une forme très normalisée de codage et/ou de structure de document (format de fichier) qui continuera d'être très largement reconnue par des systèmes informatiques sur le long terme ;
- rendre les données "autodescriptives" et "autonomes" en les "emballant" avec les métadonnées et les liens dans un logiciel qui continuera de fournir l'accès pendant un certain temps (et peut-être même "emballer" le logiciel avec les données) .
- la conversion des données dans un format où les moyens d'accès seront plus faciles à trouver ;
- la conservation des données sous leur forme originale (ou dans une version simplifiée), et la fourniture des outils qui les représenteront conformément à l'original, en utilisant soit le matériel et le logiciel d'origine (qui a également été conservé), soit en utilisant un nouveau logiciel qui émule le comportement du logiciel et/ou du matériel originaux .
- la fourniture de spécifications pour émuler les moyens d'accès originaux sur une plateforme informatique théorique intermédiaire qui puisse servir de pont pour une émulation ultérieure dans un environnement d'exploitation futur ;
- la conversion (le passage) des données dans de nouveaux formats qui sont accessibles avec chacune des nouvelles technologies d'exploitation ;
- le soutien d'un passage ultérieur en conservant les données et en enregistrant assez d'informations à leur sujet pour permettre à un utilisateur ou à un responsable futurs de les convertir sous une forme qui sera alors lisible ;
- la conservation des données et la fourniture de logiciels de nouvelle présentation (*visualiseurs*) qui restitueront une présentation acceptable de ces données pour chaque environnement d'exploitation nouveau.

### **17.12 Un soutien décisif pour les stratégies de préservation**

Quelles que soient les stratégies choisies, elles doivent être soutenues par :

- des engagements organisationnels opportuns en matière de responsabilité, de politique, de procédures et de ressources ;

- les autorisations juridiques adéquates ;
- la protection des données ;
- l'accès aux spécifications des normes et des formats de fichier pour consultation ;
- des métadonnées qui établissent l'identité, l'intégrité et les exigences techniques du matériau tout au long de sa vie ;
- une attention aux problèmes du contrôle de la qualité à toutes les étapes ;
- une surveillance des menaces telles que des changements imminents dans la technologie signifiant qu'il est nécessaire de relancer la stratégie.

### **17.13 Un plan d'éventualités**

Pour toutes les stratégies, c'est une bonne habitude que de conserver et protéger le train de données de l'objet original ainsi que les trains de données modifiés qui naîtront de la stratégie. Conserver le train de données original doit être vu comme un plan d'éventualités : il fournit une occasion de poursuivre d'autres stratégies si la stratégie retenue échoue. Une telle méthode entraîne un surcroît de dépenses pour gérer les données supplémentaires et la relation entre les trains de données parallèles. Malgré les coûts qu'elle représente, le statut incertain de la plupart des stratégies de préservation n'en rend pas moins cette méthode très attrayante.

## **LES PROBLEMES TECHNIQUES ET PRATIQUES**

### **17.14 Introduction**

Cette partie examine quelques-unes des stratégies les plus couramment proposées dans l'ordre suivant :

- les stratégies d'"investissement" (impliquant essentiellement un investissement d'effort au départ) :
  - utilisation des normes ;
  - extraction et structuration des données ;
  - encapsulation ;
  - restriction de la gamme des formats à gérer ;
  - méthode "UVC"(Universal Virtual Computer).
- les stratégies à court terme (à même de fonctionner le mieux seulement sur le court terme) :
  - technologie de préservation ;
  - compatibilité avec les versions antérieures et passage à une autre version ;
  - passage (qui peut aussi fonctionner sur des périodes plus longues).
- les stratégies du moyen au long terme (à même de fonctionner sur des périodes plus longues) :
  - (passage)
  - visualiseurs
  - émulation

- (méthode UVC)

- les stratégies de remplacement :
  - méthodes non numériques ;
  - récupération des données ;
- les stratégies conjuguées

## 17.15 Stratégies d'"investissement"

### 1. L'utilisation des normes

*Description :*

Cette stratégie implique l'utilisation de normes et de formats de fichier ouverts de préférence, largement disponibles, reconnus ou convenus, pour lesquels il existe une probabilité accrue de stabilité et un soutien à plus long terme. Ces normes ou formats peuvent soit être formellement agréés, soit être des formats normalisés *de facto* qui ont été largement adoptés par les professionnels. Une conformité aux normes peut aussi simplifier l'application ou renforcer considérablement l'efficacité des stratégies de préservation ultérieures. Cette stratégie peut se rapporter au n° 4, Réduire la gamme des formats à gérer.

Une méthode des normes particulièrement affinée est proposée concurremment avec la méthode UVC (voir ci-dessous au n° 5), c'est le *codage durable* (Gladney and Lorie, 2002). Il recommande de coder les données à valider selon les normes de traitement de données bien connues en allant jusqu'à coder les bits en format ASCII ou Unicode UTF-8, et les objets en langage XML. Pour les objets qu'il est impossible de coder ainsi, on peut appliquer cette forme de codage aux programmes qui les interpréteront et les emballer ensemble.

*Exemples :*

- Une majorité de programmes de numérisation choisissent le TIFF (Tagged Image File Format) comme norme ouverte, stable et largement reconnue pour la création d'images originales de préservation, en espérant une longue durée de vie du format.
- Le Victorian Electronic Records Strategy (VERS) stocke essentiellement les documents numériques dans l'Adobe Portable Document Format (PDF) et les encapsule dans un programme enveloppant de métadonnées XML. Le choix du PDF s'explique, en partie par la disponibilité publique de la norme exclusive, à partir de laquelle on a créé des outils indépendants de visualisation.

*Pour en savoir plus : (tous les liens ont été consultés à la date de mars 2003)*

Gibbs R., Heazlewood J. (2000), "Electronic Records - Problem Solved?: the Victorian Electronic Records Strategy and the future of electronic record keeping in Victoria". In: *Books and Bytes: Technologies for the Hybrid Library: Proceedings, 10th Biennial Conference and Exhibition, 16-18 February, 2000, Melbourne Convention Centre*. Victorian Association for Library Automation, Inc., Melbourne, 2000.

Gladney H., Lorie R. (2002). *Trustworthy 100-Year Digital Objects: Durable Encoding for When It's Too Late to Ask*. Saratoga CA, HTMG Consulting, 2002. Disponible, avec les articles ultérieurs

correspondants, auprès du HMG Consulting et via *Digital Document Quarterly*: [http://home.pacbell.net/hgladney/ddq\\_1\\_4.htm](http://home.pacbell.net/hgladney/ddq_1_4.htm)

*Quelques normes d'échange de données disponibles pour divers domaines d'activité sont répertoriées dans* : The Diffuse Project (2002). *Diffuse Standards and Specifications List*. The Diffuse Project Consortium, 2002. <http://www.diffuse.org/standards.html>

#### Les avantages potentiels de l'utilisation des normes

- Doit simplifier le processus de préservation en ralentissant le rythme des changements qui se produisent dans la technologie nécessaire à l'accès ; coder les données dans des normes très élémentaires comme l'ASCII devrait les rendre "lisibles" par système informatique pendant longtemps.
- Des formats largement reconnus peuvent avoir une gamme d'outils disponibles pour l'interprétation.
- L'utilisation de normes éditées et disponibles est plus à même de permettre la réinterprétation des données ou la reconstruction d'outils dans le futur, si nécessaire.

#### Les difficultés, les désavantages et les risques

- Peut impliquer un certain investissement pour convertir le matériau à la norme ; peut impliquer la perte de certains éléments lors de la conversion ; peut ne pas être un format normalisé disponible pour certains types d'objets.

#### Les exigences particulières

- Une connaissance des normes appropriées et une surveillance permanente de l'évolution des normes.
- Les formats de fichier normalisés doivent être bien choisis, en prenant en compte d'une part l'effet qu'a toute transformation sur les caractéristiques essentielles des objets et d'autre part l'espérance de vie des outils qui fonctionnent avec ces formats.

#### Des indications pour l'utilisation

- Il faut encourager en général l'utilisation de normes, mais plus particulièrement quand un organisme de conservation a une influence sur la création des matériaux ou sur le format dans lequel les matériaux peuvent être déposés.
- Appropriés là où ils sont ouverts, il existe des formats normalisés disponibles qui sont capables de coder dans leur complexité des objets originaux, sans perte inacceptable des caractéristiques essentielles.

#### Extraction et structuration des données

##### *Description :*

L'abstraction des données, parfois appelée aussi *normalisation*, implique d'analyser et de baliser les données de sorte que l'on puisse décrire les fonctions, les relations et la structure des éléments spécifiques qui les composent. La représentation du contenu peut être libérée des

applications spécifiques du logiciel et réalisée en utilisant différentes applications au fur et à mesure que la technologie change.

*Exemples :*

- Le San Diego Supercomputer Center a utilisé des algorithmes faits sur mesure pour appliquer des balises XML à une collection d'un million de courriels (Moore *et al.*, 2000 [2]). On a également étudié l'application de cette méthode à des documents de traitement de texte et à des ensembles de données géospatiales (Moore, 2000).

Les Archives nationales d'Australie étudient aussi actuellement cette méthode appliquée aux courriels, avec une extension ultérieure à d'autres formats (Heslop et Davis, 2001).

- Le programme du VERS du Public Record Office Victoria (Australie) étudie la représentation XML des tableaux de bases de données.
- La méthode du Universal Virtual Computer (Lorie, 2000) propose l'inclusion de balises dans les trains de données originaux pour indiquer les sections de données à interpréter en utilisant un ensemble documenté de règles pour chaque type de données.

*Pour en savoir plus :*

Heslop H., Davis S.(2002) (inédit), *An Approach to the Preservation of Digital Records*, National Archives of Australia, Canberra.

Lorie R.A. (2000), *Long-Term Archiving of Digital Information*, IBM Research Report RJ10185, IBM Research Division, San Jose, Californie. <http://domino.watson.ibm.com/library/CyberDig.nsf/7d11afdf5c7cda94852566de006b4127/be2a2b188544df2c8525690d00517082>

Moore R., Baru C., Rajasekar A., Ludaescher B., Marciano R., Wan M. *et al.* (2000), Collection-Based Persistent Digital Archives - Part 1. *D-Lib Magazine* 6(3). <http://www.dlib.org/dlib/march00/moore/03moore-pt1.html>

Moore R., Baru C., Rajasekar A., Ludaescher B., Marciano R., Wan M., *et al.* (2000), Collection-Based Persistent Digital Archives - Part 2. *D-Lib Magazine* 6(4).<http://www.dlib.org/dlib/april00/moore/04moore-pt2.html>

Moore R. (2001). Final Report for the Research Project on Application of Distributed Object Computation Testbed Technologies to Archival Preservation and Access Requirements, SDSC TR-2001-8, San Diego Supercomputer Center. <http://www.sdsc.edu/TR/TR-2001-08-.doc.pdf>

*Les avantages potentiels de l'extraction des données*

- L'indépendance de l'infrastructure simplifie le transport des données entre les plate-formes et à travers les générations de technologie.

*Les difficultés, les désavantages et les risques*

- Tous les types d'objet ne peuvent pas être extraits de cette façon.
- L'extraction exige un vaste développement des outils et des méthodes pour l'analyse et l'exploitation afin de représenter correctement et de baliser chaque type de données.

- La technologie retenue pour la présentation risque de limiter les fonctions que l'on peut représenter.

#### Les exigences particulières

- Des outils appropriés pour baliser et transformer les données.
- Un niveau élevé du contrôle de la qualité au cours du développement de méthodes pour garantir que toutes les relations et les anomalies sémantiques sont représentées.

#### Des indications pour l'utilisation

- Des données ou des documents structurés ou semi-structurés pour lesquels le maintien du contenu, de la sémantique et des relations est plus importante que des caractéristiques de présentation particulières.

### 3. L'encapsulation

#### *Description :*

L'encapsulation est un moyen largement utilisé pour lier les données et les moyens d'y accéder, de préférence dans un programme "enveloppant" qui décrit ce dont il s'agit d'une manière que peuvent comprendre une large gamme de technologies (comme un document XML). Parce qu'il est souvent peu pratique et inutile d'encapsuler les moyens réel d'accès tels que le logiciel et le matériel, l'encapsulation regroupe généralement en un seul paquet les métadonnées qui décrivent les bons outils ou qui leur sont connectées. Une autre méthode consiste à inclure, ou à relier une spécification du logiciel ou du matériel dans les métadonnées de sorte qu'elles puissent être reconstruites à l'avenir, si nécessaire.

#### *Exemples :*

- Le Reference Model for an Open Archival Information System (OAIS) décrit l'incorporation des objets de données et des métadonnées qui leur sont associées dans les Archival Information Packages (AIPs). Les métadonnées peuvent être soit regroupées directement en seul paquet avec l'objet archivé ou associées logiquement à l'intérieur du système.
- La stratégie du VERS implique la création d'"enregistrements en pelures d'oignon" dans lesquels les objets de données sont directement enveloppés dans des métadonnées codées XML, ce qui les rend indépendants d'un système de gestion.
- La stratégie du Universal Preservation Format (UPF) cherche à rendre les objets indépendants des applications ou des systèmes d'exploitation en enveloppant le contenu dans des métadonnées "autodescriptives" qui incluent les spécifications techniques permettant l'accès aux matériaux encapsulés.

#### *Pour en savoir plus*

Consultative Committee for Space Data Systems (2002). *Reference Model for an Open Archival Information System (OAIS)*. CCSDS 650.0-B-1. Blue Book. Issue 1. Janvier 2002. Washington D.C., CCSDS Secretariat, 2002. <http://www.classic.ccsds.org/documents/pdf/CCSDS-650.0-B-1.pdf>

Gibbs R., Heazlewood J. (2000). Electronic Records - Problem Solved?; the Victorian Electronic Records Strategy and the future of electronic record keeping in Victoria. In: *Books and Bytes: Technologies for the Hybrid Library: Proceedings, 10 th Biennial Conference and Exhibition, 16-18 February, 2000, Melbourne Convention Centre*. Victorian Association for Library Automation, Inc., Melbourne, 2000.

Shepard T., MacCarn D (1999). *The Universal Preservation Format: A Recommended Practice for Archiving Media and Electronic Records*. WGBH Educational Foundation, Boston. [http://info.wgbh.org/upf/pdfs/991231\\_UPF\\_RP.pdf](http://info.wgbh.org/upf/pdfs/991231_UPF_RP.pdf)

---

### Les avantages potentiels de l'encapsulation

- Produit des informations qui rendront plus facile de trouver un moyen d'accès actuel ou d'en développer un.

### Les difficultés, les désavantages et les risques

- Fournir un lien à un moyen actuel d'accès n'aborde pas vraiment le problème fondamental du changement technologique.
- Il peut être difficile de trouver ou de construire un moyen d'accès de remplacement même si les informations sont encapsulées.

### Les exigences particulières

- Une connaissance détaillée des exigences techniques pour l'accès.
- Un regroupement du paquet sûr, de sorte que les données et les métadonnées ne soient pas séparées.
- Les métadonnées décrivant les moyens de fournir l'accès doivent être tenues à jour.
- Une couche autodescriptive comme le programme d'enveloppement XML est très souhaitable.

### Des indications pour l'utilisation

- Doit certainement être considérée comme une bonne solution de base pour tous les objets qui peuvent faciliter d'autres stratégies.

#### 4. Restreindre la gamme des formats à gérer

##### *Description :*

Les programmes de préservation peuvent décider de seulement stocker les données dans une gamme limitée de formats. Cela est réalisable soit en acceptant des matériaux qui sont déjà dans ces formats, soit en convertissant les matériaux à partir d'autres formats avant le stockage.

##### *Exemples :*

- L'Archaeology Data Service (ADS) du Royaume-Uni précise une gamme privilégiée (mais pas exclusive) de formats pour le dépôt et fournit des directives pour les déposants concernant la création ou la préparation des matériaux à déposer.

*Pour en savoir plus :*

Archaeology Data Service (2001), *Guidelines for Depositors, Version 1.1*. Archaeology Data Service, York. <http://ads.ahds.ac.uk/project/userinfo/deposit.html>

*Les avantages potentiels d'une restriction de la gamme des formats*

- Réduit l'éventail des problèmes qui doivent être gérés.
- Peut servir à affiner la méthode des normes, auquel cas elle présente également les avantages de cette méthode.

*Les difficultés, les avantages et les risques*

- Ne résout pas nécessairement le problème d'accès, à moins que les formats utilisés soit efficaces lorsqu'une autre stratégie est utilisée.
- Peut restreindre la gamme des matériaux que le programme acceptera.
- La conversion risque de provoquer la perte d'éléments essentiels.

*Les exigences particulières*

- Des bases pour décider quels formats seront acceptés et comment traiter les dépôts qui ne sont pas conformes.
- Soit des règles de dépôt claires, soit un logiciel de conversion pour faire passer les données dans un autre format.
- Une vérification du contrôle de qualité rigoureuse.

*Indications pour l'utilisation*

- Des matériaux assez simples, facilement normalisés.
- Des collections avec des quantités importantes d'éléments uniformes.

5. *La méthode du "Universal Virtual Computer"*

*Description :*

La méthode du Universal Virtual Computer (UVC) cherche à spécifier une plate-forme intermédiaire, une machine virtuelle, qui est générale, mais peut être entièrement et précisément définie. Les opérations UVC sont assez simples pour être remises en oeuvre à partir des spécifications à tout moment dans l'avenir sur une plate-forme disponible.

Pour la préservation de l'objet, au moment de l'archivage on élabore un schéma logique représentant un type de données, en même temps qu'un programme de décodage qui est capable d'interpréter l'objet conformément au schéma. Le programme de décodage est écrit pour être exécuté par une mise en oeuvre de l'UVC.

Au moment de la restauration de l'objet, un émulateur adapté à l'UVC défini est mis en oeuvre sur une plate-forme disponible. L'UVC exécute le programme décodeur archivé, qui interprète l'objet archivé, et transmet les résultats à un programme restaurateur, qui restaure une représentation de l'objet conformément au schéma logique archivé.

*Exemples :*

- Le prototype "proof-of-concept" pour la méthode UVC (Lorie, 2002) a servi à créer un schéma logique, un programme décodeur et un mécanisme de représentation pour les documents PDF, tels que l'on puisse représenter le contenu du document en utilisant un programme interprète et restaurateur UVC.

*Pour en savoir plus*

Gladney H., Laurie R., (2002), *Trustworthy 10-Year Digital Objects: Durable Encoding for When It's Too Late to Ask*, Saratoga CA, HMG Consulting, 2002. Disponible, avec les articles les plus récents, auprès de HMG Consulting and via *Digital Document Quarterly*. [http://www.home.pacbell.net/hgladney/ddq\\_1\\_4.htm](http://www.home.pacbell.net/hgladney/ddq_1_4.htm)

Lorie R. (2002), *The UVC: a Method for Preserving Digital Documents - Proof of Concept*, Amsterdam, IBM Netherlands, 2002. <http://www.kb.nl/kb/ict/des/ltp/reports/4-uvc.pdf>

Lorie R. (2000), *Long-Term Archiving of Digital Information, IBM Research Report RJ10185*. IBM Research Division, San Jose, California. <http://domino.watson.ibm.com/library/CyberDig.nsf/7d11afdf5c7cda94852566de006b4127/be2a2b188544df2c8525690d00517082>

*Les avantages potentiels de la méthode UVC*

- Peut fournir des options pour préserver à la fois le comportement des matériaux document-type et celui des programmes logiciels.
- Une plate-forme intermédiaire définie, unique, peut réduire le travail de développement nécessaire pour loger des combinaisons différentes de logiciel et de plate-forme.
- Les spécifications de l'UVC sont conçues pour être simples, permettant une utilisation par des programmeurs de compétence moyenne et simplifiant la construction d'interprètes et d'émulateurs UVC dans l'avenir.
- Peut être conçue pour interpréter le train de données de l'objet original, ou une représentation transformée ou abstraite.
- Les codages de données et les programmes décodeurs peuvent être testés au moment de leur création sur une installation UVC contemporaine. On pourra alors s'attendre que les futures installations de spécifications UVC reproduisent le comportement actuel.

*Les difficultés, les désavantages et les risques*

- La méthode est actuellement en développement et a été élaborée sous forme de prototype pour une représentation transformée d'un document original. Elle nécessite un travail supplémentaire pour être appliquée à des programmes logiciels. Comme avec l'émulation (voir n° 11), la complexité des comportements du programme peut poser des problèmes.

- L'investissement nécessaire au moment de l'archivage dans le développement de méthodes de codage ou de programmes interprètes adaptés UVC pour chaque type de données.
- Peut exiger un soutien important de la part des producteurs de l'information en vue de fournir des versions compatibles UVC de leurs produits (Gladney and Lorie, 2002) pour lesquels ils risquent d'avoir peu d'occasions promotionnelles ou commerciales.
- L'investissement nécessaire au moment de la restauration dans le développement d'un émulateur UVC et de programmes de restauration.
- Si les objets originaux des données sont abstraits ou transformés à des fins de codage, ces transformations risquent de mettre au rebut les caractéristiques essentielles.

#### Les exigences particulières

- Le développement d'un schéma logique ou d'une représentation pour chaque type de données ou programme au stade du codage.
- Le développement au stade du codage d'un programme décodeur pour interpréter chaque type de données ou de programme, écrit pour l'exécution par un interprète ou un émulateur UVC.
- Le développement d'un interprète ou d'un émulateur UVC au moment de la restauration de l'objet, pour s'adapter à une plate-forme répandue.
- Le développement de programmes de restauration pour renvoyer une représentation de l'objet original, fondé sur le schéma logique et les données récupérées par l'UVC en exécutant le programme décodeur archivé.
- L'archivage de l'objet de données ou du programme (ou de sa représentation transformée), de tous les schémas logiques associés, du programme décodeur exécutable-UVC, et de la spécification UVC et des instructions de restauration.
- Une compétence suffisante pour le développement des schémas logiques, le codage, les programmes décodeurs, la mise en application de l'émulateur UVC d'après la spécification et les programmes de restauration.

#### Des indications pour l'utilisation

- Peut convenir là où les objets sont suffisamment représentés, codés, interprétés et restaurés en utilisant des outils développés d'après la spécification. Au moment où nous écrivons ces lignes, la spécification UVC est en cours de développement.

### **17.16 Les stratégies à court terme**

#### 6. La préservation de la technologie

##### *Description :*

Cette stratégie implique de conserver et d'entretenir le logiciel et le matériel originaux avec lesquels les objets numériques ont été présentés. C'est la première démarche fondamentale, et d'une certaine manière, la plus importante dans la préservation de l'accès si aucune autre stratégie n'est en place. Si le matériel et le logiciel nécessaires pour l'accès sont mis au rebut avant que d'autres

stratégies soient disponibles, il peut devenir effectivement impossible de fournir un accès ultérieur sans un travail de récupération des données coûteux et aléatoire.

*Exemples :*

- Entretenir les vieux lecteurs de disquette qui accepteront des disquettes dont le format n'est plus accepté par l'équipement informatique actuel.
- Entretenir le logiciel dépassé pour l'utiliser avec les matériaux qu'il a légués.
- Entretenir les vieux systèmes d'exploitation pour permettre l'utilisation du logiciel qui ne fonctionne pas sur les plate-formes actuelles.

*Pour en savoir plus*

Jones M., Beagrie N. (2001), *Preservation Management of Digital Materials: A Handbook*, The British Library, Londres, 2001.

*Les avantages potentiels de la préservation de la technologie*

- Présenter les objets numériques au moyen du logiciel et du matériel qui leur étaient destinés à l'origine garantit que toute la gamme des éléments et des fonctions voulus sera présentée.
- Accorde un délai pendant lequel d'autres stratégies peuvent être développées ou mises en oeuvre.
- Comme avantage secondaire, documenter le matériel et le logiciel qu'il faut conserver peut mener à une meilleure compréhension de la collection et de ses annexes, permettant ainsi d'apporter des informations utiles à la planification et à la mise en oeuvre d'autres stratégies.

*Les difficultés, les désavantages et les risques*

- Assurer l'entretien à long terme de l'équipement, avec des pièces et des compétences de plus en plus difficiles à trouver, est très peu probable.
- Même avec une gestion efficace, la fenêtre d'accès utilisant cette méthode peut ne pas dépasser cinq à dix ans à partir du moment où le format original est périmé.(Mais cela vaut peut-être mieux que de perdre immédiatement l'accès.)
- Nécessite la gestion et l'entretien d'une large gamme d'équipement et de logiciels, en même temps que celle de matériaux complémentaires comme les manuels et les licences, qui peuvent être difficiles et coûteux à réaliser.
- Les compétences et l'assistance technique nécessaires risquent tout simplement ne pas être disponibles.

*Les exigences particulières*

- Exige une identification active du matériel et du logiciel nécessaires pour l'accès.
- Exige des dispositifs d'entretien actifs et permanents pour l'équipement et des dispositifs de préservation et de licence pour le logiciel.

- Exige des mesures pour garantir que les compétences sont partagées et ne dépendent pas d'une seule personne.
- Un certain nombre d'organismes peuvent regrouper des équipements ou des pièces dépassés, et utiliser des archives de logiciel utilisées en commun ou par tiers.
- Par principe, si le logiciel d'accès nécessaire est disponible, il faut se le procurer et le conserver au moins jusqu'à ce qu'on ait mis en place une autre stratégie.
- Le logiciel conservé doit être traité comme tous les autres objets numériques : il réclame un contrôle, une documentation, un rafraîchissement et un entretien des supports, et il est soumis aux contraintes d'un copyright.

#### Des indications pour l'utilisation

- Recommandée comme stratégie de départ pour tous les programmes de préservation, en l'absence de stratégies à long terme ou pendant qu'elles sont en cours d'élaboration.
- Peut être la seule option disponible pendant une période plus longue pour les objets numériques complexes comme les objets logiciels et multimédia.
- Recommandée pour les logiciels qui doivent assurer le suivi d'une gamme d'autres stratégies.

#### 7. La compatibilité en amont et le passage à une autre version

##### *Description :*

Cette stratégie s'appuie sur la capacité qu'ont certains logiciels d'interpréter et de présenter des objets créés avec les versions antérieures du même logiciel. Dans le cas de la compatibilité en amont, la présentation peut se limiter à une visualisation temporaire ; en revanche, le passage à une autre version convertit en permanence les documents dans un format que la version actuelle du logiciel peut présenter.

##### *Exemples :*

- Les logiciels de navigation sur la Toile sont généralement capables d'interpréter et de présenter des matériaux écrits en utilisant les versions antérieures de la norme HTML.
- Les applications de documents de bureau, comme le traitement de texte, l'application de calcul ou de base de données, permettent généralement de transformer et de réenregistrer les versions antérieures de leurs formats de fichier en une nouvelle version, dans le cadre du parcours de modernisation de l'application.
- Le projet Digital Preservation Testbed (Digitale Testbed Bewaring), aux Pays-Bas, a étudié le passage à une autre version à travers et sur des générations de versions d'application.

Pour en savoir plus

Potter M. (2002), Researching Long Term Digital Preservation Approaches in the Dutch Digital Preservation Testbed (Testbed Digitale Bewaring), *RLG DigiNews* 6(3). <http://www.rlg.org/preserv/diginews/v6-n3-a2.html>

### Les avantages potentiels de la compatibilité en amont

- La disponibilité : les développeurs de logiciels intègrent souvent une compatibilité en amont appropriée ou un chemin de passage à une autre version pour les documents.
- Peut prolonger le délai avant qu'une transformation ou un traitement plus étendus soient nécessaires.
- Dans certains cas, offre une fonctionnalité similaire à la présentation originale.

### Les difficultés, les désavantages et les risques

- Il est improbable que la compatibilité soit maintenue sur des générations de logiciels.
- Susceptible d'introduire des modifications indésirables de façon incrémentielle si elle est utilisée sur de nombreuses générations.
- Ces chemins risquent de ne pas être disponibles pour tous les types d'objets.
- Peut être abandonnée par les développeurs de logiciels pour une nouvelle génération de leur logiciel, aussi la sûreté de fonctionnement est-elle imprévisible.
- Même entre les versions les plus proches des mêmes applications, il peut y avoir des modifications indésirables introduites dans les matériaux.

### Les exigences particulières

- Comme dans toutes les opérations de passage, il est indispensable qu'il y ait une vérification du contrôle de la qualité pour détecter tous les changements inacceptables.

### Des indications pour l'utilisation

- Peut fournir un itinéraire de passage simple et à court terme pour des objets types document dans des formats qui proposent une succession de versions, tant que les conversions n'introduisent pas de modifications indésirables.
- Peut remplacer la préservation de la technologie pour des objets comme les feuilles de calcul et les bases de données complexes, pour lesquelles aucune autre stratégie n'est encore disponible.

## 8. Le passage

### *Description :*

Le passage consiste dans le transfert des matériaux numériques d'une génération de matériel ou de logiciel à une autre. Distinct du rafraîchissement, qui conserve le train de données en le transférant d'un support à un autre, le passage entraîne la transformation de la forme logique d'un objet numérique, de sorte que l'objet conceptuel peut être restitué ou présenté par un nouveau matériel ou logiciel.

Il y a un certain nombre de stratégies que l'on peut considérer comme des formes de passages : elles diffèrent par le moment où survient la transformation et par les types d'objets transformés. La méthode de passage la plus couramment proposée consiste à transformer en

permanence un format logique dans un autre, conformément au changement technologique, si bien que tous les objets ainsi transférés peuvent être présentés avec la technologie actuelle.

Il est également possible de proposer un modèle de "passage à la demande" ou "de passage au point d'accès". Cette méthode est étudiée au point n° 10 ("Les visualiseurs") ci-dessous.

NB. En raison de l'effets cumulatif probable des passages répétés, cette méthode a été incluse dans les stratégies à court terme. Toutefois, il est probable que pour certains types de données et de format, ce passage se révèle être une stratégie utile à long terme.

*Exemples :*

- Des collections de matériaux hétérogènes aux formats bien définis, comme des collections d'images, peuvent être adaptées à une transformation de format.

*Pour en savoir plus :*

Lawrence G.W, Kehoe W.R., Rieger O.Y., Walters W.H., Kenney A.R. (2000), *Risk Management of Digital Information: A File Format Investigation*, Council on Library and Information Resources, Washington, D.C. <http://www.clir.org/pubs/reports/pubs93/contents.html>

National Archives of Australia, *Managing Electronic Records*

*Appendix 3: Preserving Electronic Records through Migration*. National Archives of Australia, Canberra. [http://www.naa.gov.au/recordkeeping/er/manage\\_er/append\\_3.html](http://www.naa.gov.au/recordkeeping/er/manage_er/append_3.html)

Task Force on Archiving of Digital Information (1996), *Preserving Digital Information: Report of the Task Force on Archiving of Digital*

*Information*, Commission on Preservation and Access and Research Libraries Group. <ftp://ftp.rlg.org/pub/archtf/final-report.pdf>

*Les avantages potentiels du passage*

- Des procédures simples de passage sont bien établies pour certains formats.
- Les passages effectués pour répondre aux changements de technologie permettent de vérifier les objets transférés par rapport aux copies non transférées pour voir si les éléments essentiels ont bien été conservés.
- Si le passage s'est opéré comme il convient, les utilisateurs peuvent s'attendre en toute confiance que le matériau sera présenté avec la technologie régnante, sans avoir besoin d'un matériel ou d'un logiciel particuliers.

*Les difficultés, les désavantages et les risques*

- On risque de se trouver dans l'impossibilité de fournir l'accès à certains matériaux comme les objets complexes utilisant le transfert de format, faute d'avoir le moyen de représenter des fonctions complexes dans le nouveau format.
- Même si elle semble fonctionner, la transformation du codage logique risque de compromettre l'intégrité des éléments essentiels du matériau.

- Les objets devront être transformés régulièrement pour suivre la technologie, créant ainsi une charge de frais permanents. Les passages à grande échelle signifient une analyse détaillée des structures de données, l'élaboration de règles pour contrôler la transformation, une écriture de programmes pour modifier le codage des données, ainsi qu'un contrôle de qualité et un "nettoyage" approfondis. Cela se justifie facilement pour d'importantes bases de données commerciales sensibles, mais ces mesures rigoureuses risquent d'être irréalisables quand il s'agit de matériaux moins sensibles dans une gamme diverse de formats de fichier.
- Les petites modifications qui surviennent entre générations peuvent, en s'accumulant, aboutir à des changements importants ou à des pertes dues à des passages répétés.

#### Les exigences particulières

- Exige des programmes et des outils pour effectuer la conversion.
- Une vérification rigoureuse du contrôle de la qualité, à la fois au cours du développement des méthodes et après le passage.
- La documentation de la méthode de passage doit être stockée dans les métadonnées, comme faisant partie de l'histoire de l'objet et de son authenticité.
- Les processus de passage devront, si possible, être entièrement réversibles, en documentant la nature et l'emplacement de toutes les transformations.
- Autre solution : conserver une copie des objets numériques d'origine si une transformation n'est pas réversible ou si certains éléments essentiels risquent d'être perdus. (Conserver une copie du format original est, de toute manière, une excellente pratique.)
- Le processus de passage doit être testé avant d'être définitivement mis en oeuvre, et sa réussite avérée avant de détruire toutes les générations intermédiaires.

#### Des indications pour l'utilisation

- Le passage peut être adapté à toute une gamme d'objets numériques, en particulier les types d'objets documents et bases de données.
- Là où les éléments essentiels qu'il faut préserver sont assez simples et ne dépendent pas de la présentation du matériau et n'impliquent pas des fichiers exécutables.
- Peut être très compétitive pour des collections homogènes comme les collections d'images et d'audio numériques qui sont dans des formats bien normalisés, très répandus et non exclusifs.
- Certains formats exclusifs et très répandus peuvent aussi convenir si les propriétaires de brevet et de licence acceptent de les fournir ou autorisent d'autres usagers à développer des spécifications de format ou des outils de conversion.

## 9. Le réagencement

### *Description :*

Les objets logiciels qui dépendent fortement d'un système ou d'une plate-forme spécifiques pour pouvoir fonctionner sont peut-être les plus touchés par les changements de technologie ; ils sont aussi en général inadaptés à de nombreuses stratégies de préservation, y compris des passages réguliers. Le réagencement de logiciels offre plusieurs stratégies pour transformer les logiciels au fur et à mesure que les technologies changent, similaires à la transformation des formats de données. Entre autres possibilités :

- la mise au point et la recompilation du code d'origine pour une plate-forme nouvelle ;
- la rétrotechnique du code compilé dans un code de niveau supérieur et en l'adaptant à la nouvelle plate-forme ;
- le recodage du logiciel à partir de zéro, ou le recodage dans un autre langage de programmation (Wheatley, 2001) ;
- la traduction d'instructions binaires compilées pour une plate-forme directement en instructions binaires pour une autre plate-forme. (Des chercheurs de l'Université de Queensland (Cifuentes *et al.*, 1999) étudient ce concept.)

### *Pour en savoir plus*

Cifuentes C., Van Emmerik M., Ramsey N., (1999). The Design of a Resourceable and Retargetable Binary Translator. In: *Proceedings: Sixth Working Conference on Reverse Engineering, October 6-8, 1999, Atlanta, Georgia, USA*. IEEE Computer Society, New Jersey, 1999, p. 280-291.

Wheatley P. (2001). Migration - a CAMiLEON discussion paper, *Ariadne* 29. <http://www.ariadne.ac.uk/issue29/camileon/>

### Les avantages potentiels du réagencement

- Offre le potentiel de transférer les objets logiciels d'une plate-forme à une autre.

### Les difficultés, les désavantages et les risques

- Sauf pour les programmes au code source libre et pour les logiciels développés en interne, le code source est souvent indisponible ou soumis à des droits d'utilisation.
- Même quand le code source est disponible, le transfert à d'autres plate-formes n'est pas insignifiant, et en général, la nouvelle plate-forme a besoin de programmes compilateurs et interprètes pour le langage source.
- Nécessite de fournir un temps et un effort considérables pour chaque objet.
- Toute forme de réagencement est généralement strictement interdite par les accords de licence de l'utilisateur normal et enfreint gravement les droits de propriété intellectuelle. D'autres formes de transformation peuvent aussi violer ces droits.

### Les exigences particulières

- Un haut niveau de compétence.
- Des outils pour transformer un code lisible par l'homme en un code lisible par la machine.
- L'autorisation expresse de réagencer.

### Des indications pour l'utilisation

- Doit être envisagée uniquement là où les droits sont formellement accordés, et quand la compétence, les outils et, de préférence, le code source sont disponibles.

## **17.17 Les stratégies de moyen à long terme**

### 10. Les visualiseurs et le passage au point d'accès

#### *Description :*

On a proposé un certain nombre de solutions de rechange au passage périodique et incrémentiel. Elles consistent en l'utilisation de visualiseurs, d'outils logiciels ou de méthodes de transformation qui fournissent l'accessibilité au moment de l'accès et utilisent le train de données original.

#### *Exemples :*

- La méthode de "passage à la demande" développée conjointement avec les projets CEDARS et CAMiLEON inclut un outil logiciel dans les objets numériques et utilise les métadonnées de l'objet pour enregistrer une méthode qui permet d'accéder à l'objet au moyen de l'outil. Au fur et à mesure que la technologie change, les métadonnées sont actualisées pour refléter les modifications dans la méthode d'accès (Cedars, 2002 ; Mellor Sergeant and Wheatley, 2003).
- La méthode TOMS (Typed Object Model Server) fournit des méthodes de transformation pour les types de documents et de données courants, permettant à un serveur de choisir un parcours de transformation adapté à une gamme de types d'objet (Thibodeau, 2002).
- La stratégie VERS convertit les documents dans un format PDF selon le principe que l'on peut construire des visualiseurs tiers pour PDF à partir de la spécification du format.

La méthode Rosetta Stones comporte des règles pour l'interprétation du format des données et des fichiers échantillons à la fois dans le format original et dans un format de référence qui montre à quoi les fichiers ressembleraient s'ils étaient correctement interprétés. Les outils logiciels peuvent alors être construits pour suivre la méthode d'interprétation proposée pour les fichiers, et pour vérifier si l'interprétation est correcte en comparant les fichiers échantillons à la présentation de référence (Thibodeau, 2002).

#### *Pour en savoir plus :*

Cedars Project (2001), *The Cedars Project Report, April 1998 - March 2001*, Cedars, Université de Leeds. <http://www.leed.ac.uk/cedars/pubcom/papers/projectReports/Cedars>

[ProjectReportToMar01.pdf](#)

Cedars Project (2002), *Cedars Guide to: Digital Preservation Strategies*, Cedars, University of Leeds. <http://www.leeds.ac.uk/cedars/guideto/dpstrategies/dpstrategies.html>

Mellor P., Seargent D., Wheatley P. (2002), *Migration on Request; A Practical Technique for Preservation*, CAMiLEON Project, Université du Michigan. <http://www.si.umich.edu/CAMILEON/reports/migreq.pdf>

Thibodeau K. (2002), Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years, in *The State of Digital Preservation: An International Perspective - Conference Proceedings, Documentation Abstracts, Inc., Institutes for Information Science, Washington, D.C., April 24-25, 2002*, Council on Library and Information Resources, Washington, D.C. <http://www.clir.org/pubs/reports/pub107/thibodeau.html>

#### Les avantages potentiels de l'emploi de visualiseurs, etc.

- Le train de données original est interprété et présenté par le visualiseur, des outils ou la méthode de transformation plutôt que par un train de données transféré de manière incrémentielle, aussi peut-on éviter le risque de déformations cumulatives du contenu ou de la fonction sur des générations de passages.
- Les objets sont interprétés ou transformés uniquement lorsqu'on y accède, de sorte que l'on évite le coût représenté par un passage régulier des objets sans se soucier de la demande d'accès.

#### Les difficultés, les désavantages et les risques

- Il peut ne pas y avoir de visualiseurs ou d'outils disponibles pour des matériaux complexes, y compris les fichiers exécutables.
- Les visualiseurs peuvent être capables de représenter certains (mais pas tous) des éléments de certains matériaux (mais cela peut être un avantage là où la fonction "visualisation seulement" est nécessaire.
- Il risque d'y avoir un écart trop grand entre le format original et les technologies en usage au moment de l'accès pour que les outils ou les méthodes le surmontent.
- Les visualiseurs, les outils ou les méthodes, et les métadonnées correspondantes doivent aussi être conservés ou ajustés au fur et à mesure que les technologies changent.

S'ils ne sont pas essayés au préalable, les visualiseurs, les outils ou les méthodes risquent de ne pas pouvoir présenter les objets conceptuels d'une façon satisfaisante.

#### Les exigences particulières

- Toute la documentation des formats de fichier et des méthodes de transformation doit être tenue à jour.
- Un maintien à niveau rigoureux des métadonnées techniques pour parer aux changements technologiques.
- Les métadonnées techniques et les méthodes d'accès doivent être liées mais stockées séparément des objets numériques, de sorte que les métadonnées ou les méthodes puissent être actualisées centralement.

### Des indications d'utilisation

- Peut être préférée au passage périodique là où les coûts impliqués par des passages répétés posent un problème ou lorsqu'il peut y avoir de longs intervalles entre les demandes d'accès.
- Peut convenir lorsqu'on démontre par avance qu'il est réalisable de construire des outils ou des visualiseurs qui interpréteront les formats de fichier à partir des instructions, des spécifications ou des méthodes incluses.

## 11. L'émulation

### *Description :*

L'émulation nécessite l'emploi d'un logiciel qui fait se comporter une technologie comme une autre. Dans le contexte de la préservation numérique à long terme, cela impliquera de faire se comporter les futures technologies comme l'environnement original d'un objet numérique préservé, de sorte que l'objet original puisse être présenté sous sa forme originale à partir du train de données original.

L'émulation matérielle est souvent proposée comme une stratégie largement applicable, car les spécifications du matériel sont susceptibles d'être plus complètes ou plus facilement définies que celles du logiciel. L'émulation d'une plate-forme matérielle peut se révéler également fort utile dans la mesure où elle permet à une gamme de systèmes et d'objets numériques de fonctionner, résolvant ainsi le problème pour une gamme très étendue d'objets numériques. Une autre possibilité consiste à envisager l'émulation d'applications spécifiques logicielles ou de comportements. Un argument contre ceci, c'est que chaque application exigera des efforts d'émulation individuels. d'un autre côté, si les besoins d'émulation sont faibles, il peut être superflu de dépenser des efforts dans l'émulation d'une plate-forme ou d'un système entiers pour un petit nombre ou une petite gamme d'objets.

### *Exemples :*

- Des chercheurs travaillant pour le projet CAMiLEON ont étudié l'émulation en tant que méthode de préservation numérique, y compris l'utilisation expérimentale d'émulateurs disponibles (Hedstrom and Lampe, 2001) et la construction d'un émulateur pour un système de 1970, George3 (Holdsworth and Wheatley, 2001).
- Un Universal Virtual Computer (UVC) a été proposé comme une plate-forme intermédiaire qui servirait sur des systèmes futurs, de sorte que l'émulation de comportement de programme ait pour cible une seule plate-forme permanente, réduisant ainsi au maximum les couches supplémentaires d'émulation (Lorie, 2000).
- La possibilité de différer la construction de l'émulateur jusqu'à ce qu'il soit nécessaire, en préservant à la place des spécifications détaillées pour des émulateurs qui seraient créés quand on en aurait besoin, a également été proposée (Rothenburg, 2000).

### *Pour en savoir plus*

Hedstrom M., Lampe C. (2001), Emulation vs. Migration: Do Users Care? *RLG DigiNews* 5(6). <http://www.rlg.org/preserv/diginews/diginews5-6.html#feature1>

Holdsworth D., Wheatley P. (2001), Emulation, Preservation and Abstraction, *RLG DigiNews* 5(4). <http://www.rlg.org/preserv/diginews/diginews5-4.html#feature2>

Lorie RA (2000), *Long-Term Archiving of Digital Information*, IBM Research Report RJ10185, IBM Research Division, San Jose, Californie. <http://domino.watson.ibm.com/library/CyberDig.nsf/7d1afdf5c7cda94852566de006b4127/be2a2b188544df2c8525690d00517082>

Rothenberg J. (2000), *Using Emulation to Preserve Digital Documents*, Koninklijke Bibliotheek, La Haye. <http://www.kb.nl/kb/pr/fonds/emulation/usingemulation.pdf>

### Les avantages potentiels de l'émulation

- L'émulation est un principe établi de la science informatique, et elle est souvent utilisée pour développer et tester les nouveaux logiciels avant la fabrication.
- Il existe aujourd'hui des émulateurs pour divers systèmes et plate-formes, qui vont de l'émulateur pour systèmes obsolètes construits par des passionnés, à des systèmes commerciaux pour l'utilisation de multiplate-formes ou pour tester les logiciels.
- Dans sa plus large application possible, l'émulation permettra de recréer une gamme d'objets numériques avec toutes leurs fonctions, y compris des objets logiciels, en utilisant le train de données original, non transformé, en combinaison avec le logiciel original préservé.

### Les difficultés, les désavantages et les risques

- L'émulation est techniquement complexe ; elle exige un effort soutenu et une compétence spécifique, aussi peut-elle être très coûteuse.
- En tant que technique de préservation numérique très largement applicable, l'émulation en est encore au stade de la recherche.
- L'efficacité de l'émulation peut être contrecarrée par une documentation inadéquate du logiciel, ou par une utilisation non normalisée de formats de fichier comme solution improvisée.
- Au fur et à mesure que les systèmes deviennent plus complexes, les demandes d'émulation, laquelle aura à intégrer des composants multiples, le deviennent également. L'émulation de tous les aspects d'un système ou d'une application peut se révéler impossible.
- Les futurs utilisateurs auront peut-être du mal à savoir comment réagir à une large gamme d'applications archaïques fonctionnant sous émulation. Aussi, des outils de présentation contemporains seront-ils probablement toujours nécessaires, ajoutant une couche supplémentaire d'outils logiciels évolutifs qui sont nécessaires pour accéder à l'objet émulé.
- Comme la technologie et les plate-formes évoluent avec le temps, les émulateurs eux-mêmes devront, soit passer à une nouvelle plate-forme, ce qui conduit en puissance à des couches sur des couches d'émulateurs, soit faire émuler leurs ordinateurs centraux.

### Les exigences particulières

- Un niveau de compétences suffisant pour développer des émulateurs, ou accéder à des émulateurs développés par quelqu'un d'autre.
- Une documentation complète et précise des systèmes à émuler.
- Une grande clarté sur le niveau de l'émulation requis, par exemple l'émulation de tout le matériel contre des fonctions logicielles spécifiques.

- Il faut créer un code d'émulation qui utilise les techniques normalisées du génie logiciel, notamment une bonne structure de code ainsi qu'un commentaire et une documentation approfondis.
- Le code du programme d'émulation devra être écrit en code source libre, dans un langage de programmation normalisé avec de bonnes perspectives de longévité et de compatibilité future.
- Tout code nécessaire qui n'est pas normalisé (par exemple pour des fonctions périphériques spécifiques) devra être écrit comme un module séparé mais relié, et bien documenté.

#### Des indications pour l'utilisation

- Là où des émulateurs adaptés sont déjà disponibles pour les plate-formes requises.
- Là où des compétences suffisantes sont disponibles pour la construction d'un émulateur.
- Pour des objets très complexes ou pour certains objets comme les logiciels exécutables, qui ne peuvent fonctionner qu'avec des systèmes ou du matériel spécifiques.
- Pour des objets dont la valeur réside dans le fait d'être visualisés dans leur environnement original.

#### **17.18D'autres stratégies de rechange**

---

#### 12. Méthodes non numériques

##### *Description :*

Une méthode de remplacement des méthodes de préservation numérique consiste à "imprimer" les objets sur des supports analogiques assez stables, comme du papier, un microfilm ou des plaques de nickel (comme dans la technologie HD-Rosetta, qui micro-grave des images document sur du nickel avec un faisceau ionique et permet de le visualiser avec des agrandisseurs optiques), déplaçant ainsi la charge de la préservation d'un support à un autre en la confiant à une copie analogique au lieu d'un objet numérique.

##### *Exemples :*

- Une institution a la garde d'anciennes bases de données dans un format protégé qui est périmé et qui sera illisible lors de la prochaine modernisation du système sans qu'il y ait actuellement aucun moyen d'extraire ou de transférer sur d'autres supports les données qu'elles contiennent. L'institution choisit d'imprimer la totalité du contenu des données sur papier comme autant d'enregistrements particuliers pour préserver au moins un accès aux données, même s'il est inefficace. L'institution conserve aussi les bases de données sous forme numérique, au cas où un mécanisme d'accès deviendrait disponible.
- Un programme de numérisation crée un microfilm de sortie d'ordinateur à partir d'originaux d'images numériques en tant que sauvegarde physique de la collection et autre source de préservation, de distribution et d'accès.

*Pour en savoir plus :*

Hedstrom M., Lampe C. (2001), Emulation c. Migration: Do Users Care? *RLG DigiNews* 5(6). <http://www.rlg.org/preserv/diginews5-6.html#feature1>

Norsam Technologies (2001), *HD-Rosetta Archival Preservation Services*. <http://www.norsam.com/hdrosetta.htm>

#### Les avantages potentiels des méthodes non numériques

- Les objets sont saisis sous une forme lisible par l'homme et échappent à la menace d'une obsolescence technologique et à la pression exercée par les cycles permanents de préservation numérique.
- Fournit une méthode de préservation plus simple, car on peut préserver les matériaux analogiques sur le long terme en utilisant des méthodes de préservation traditionnelles.
- Susceptible d'impliquer un coût de conversion pour une seule fois.

#### Les difficultés, les désavantages et les risques

- Risque de perdre les avantages offerts par la technologie numérique comme la commodité d'utilisation ou l'efficacité de stockage.
- Perd les fonctions caractéristiques qui reposent sur la technologie numérique, comme les calculs sur tableur, le son incorporé ou les images mobiles (même si certaines d'entre elles peuvent aussi être sauvegardées sur des formes analogiques séparées), les fonctions de recherche et de navigation. (Cela peut ne pas être un désavantage si ces éléments n'ont pas besoin d'être préservés.)
- N'échappe pas entièrement au risque d'une obsolescence technologique puisque la technologie de lecture pour certains formats peut évoluer avec le temps. Il est toujours possible en théorie d'utiliser des agrandisseurs optiques pour lire les informations, mais cela peut être peu pratique, en rendant certains matériaux effectivement inutilisables pour la plupart des utilisateurs. Pour les enregistrements sonores, l'accès analogique risque d'être moins fiable que l'accès numérique.
- La stabilité à long terme des supports analogiques dépendra d'environnements de stockage coûteux qui se révèlent être moins fiables que des systèmes informatiques bien gérés fondés sur un niveau élevé de redondance.

#### Les exigences particulières

- Les matériaux du support utilisés pour la conversion en analogique doivent présenter une qualité d'archivage et être stockés dans des conditions d'archivage.
- Il est toujours recommandé de conserver les objets numériques, dans la mesure du possible, au cas où on développerait dans le futur un chemin d'accès approprié.

#### Des indications pour l'utilisation

- Seulement adaptée à des objets qui ne réclament pas les fonctions de la technologie numérique pour parvenir à leurs fins, par exemple des documents type texte, image ou données qui ne réclament aucune fonction au-dessus de l'affichage "plat".

- Peut être une mesure pragmatique en attendant que d'autres stratégies soient développées.
- Peut être nécessaire en dernier ressort là où aucune stratégie n'est disponible, une accessibilité limitée étant préférable à pas d'accessibilité du tout.

### 13. La récupération des données

#### *Description :*

La récupération des données (parfois appelée *archéologie des données*) consiste habituellement à récupérer d'abord des données sous forme de bits à partir des supports physiques puis, au terme d'une série d'étapes, à restaurer l'intelligibilité des données ainsi récupérées. On s'en sert le plus souvent pour récupérer des données à partir de supports défectueux ou détériorés, mais des méthodes pour restaurer l'intelligibilité ont aussi été utilisées pour sauver des documents dans des formats périmés. Toutefois, prétendre que l'on sera capable de réaliser un tel sauvetage dans le futur est une solution hasardeuse et à très haut risque qui ne saurait remplacer un programme de préservation mis en route dès à présent.

#### *Exemples :*

Le UK Archaeology Data Service a réalisé une récupération de données sur des disques provenant du Newham Museum Archaeological Service (Dunning, 2001.) On avait découvert qu'un certain nombre de fichiers étaient détériorés et irrécupérables. Un grand nombre de ceux que l'on a récupérés existaient dans des formats de données périmés qui nécessitaient un logiciel spécialisé pour être interprétés, ou étaient insuffisamment documentés, de sorte qu'on ne pouvait pas établir de façon satisfaisante le contexte des données.

#### *Pour en savoir plus :*

Ross S., Gow A. (1999), *Digital Archaeology: Rescuing Neglected and Damaged Data Resources*, Library Information Technology Centre, Université de South Bank, Londres. <http://www.ukoln.ac.uk/services/elib/papers/supporting/pdf/p2con.pdf>

Woodyart D. (2001). Data Recovery and Providing Access to Digital Manuscripts. Article présenté à Digital Dancing: New Steps, New Partners - Information Online 2001, Tenth Exhibition and Conference, 16-18 janvier 2001. Sydney Convention and Exhibition Centre, Darling Harbour, Sydney. [http://www.nla.gov.au/nla..arc-14099-20020211-www.csu.edu.au/speciam/online2001/papers/digital\\_issues\\_ii.htm](http://www.nla.gov.au/nla..arc-14099-20020211-www.csu.edu.au/speciam/online2001/papers/digital_issues_ii.htm)

Dunning A.(2001), *Excavating Data - The Retrieval of the Newham Archive*, Arts and Humanities Data Service. <http://ahds.ac.uk/newham.pdf>

#### Les avantages potentiels de la récupération de données

- Peut permettre la récupération de données qui autrement seraient définitivement perdues.

#### Les difficultés, les désavantages et les risques

- Il n'y a aucune garantie de récupération à partir des supports, ni de récupération de l'intelligibilité des données.
- Sans une documentation suffisante, l'interprétation des données est souvent plus qu'approximative et l'identité, l'intégrité et le contexte sont difficiles à établir.

- Souvent coûteuse ; exige aussi des efforts importants par élément.
- Sans une documentation suffisante, il est impossible d'estimer à l'avance si les efforts et la dépense seront justifiés.

#### Les exigences particulières

- Très aidée par une bonne documentation des types de fichier et du contenu.
- Peut nécessiter les services judiciaires spécialisés dans la récupération de données ou un logiciel de reconnaissance.

#### Des indications pour l'utilisation

- Utilisation recommandée seulement en tant que stratégie de récupération et de restauration de données en cas de détérioration des supports, ou lorsque l'on découvre des supports ou des formats de fichier périmés et lorsque la valeur ou l'importance des données sont à même de justifier les coûts potentiels.

### **17.19 Les stratégies conjuguées**

Comme on l'a précédemment noté, il peut être nécessaire de disposer, pour une collection diverse, d'un certain nombre de stratégies qui embrassent toute la gamme des objets et des caractéristiques à préserver ; différentes méthodes peuvent également bien s'articuler entre elles avec le temps. Les programmes de préservation doivent aussi prendre en compte les avantages potentiels de la redondance en poursuivant plus d'une stratégie : même avec une bonne planification, une stratégie unique risque d'échouer, laissant le programme sans moyen d'accès. Plusieurs exemples cités ci-dessus ne se contentent pas d'une méthode ; par exemple :

- des normes comme le TIFF pour les collections d'images sont souvent choisies pour préparer un passage à d'autres formats normalisés sur le long terme ;
- la stratégie du VERS marie l'utilisation de normes (PDF, XML) pour l'utilisation future de visualiseurs à l'éventuel passage à des métadonnées codées dans l'avenir ;
- les archives permanentes (Moore, 2001) utilisent l'abstraction des données en vue d'un éventuel passage - passage des données, du système de balisage et du logiciel de soutien, et la modernisation du matériel ;
- la méthode UVC (Universal Virtual Computer) conjugue l'abstraction des données avec des règles pour le passage des objets de données au point d'accès, et une méthode d'émulation pour les objets logiciels. La méthode du "codage durable" ajoute l'utilisation de normes fondamentales pour le codage des données, y compris le codage de ce qui pourrait être compris par l'UVC.

## POUR LES PROGRAMMES DE PRESERVATION DISPOSANT DE PEU DE MOYENS

### 17.20 Choisir des options à bas coûts

Les programmes de préservation disposant de peu de ressources risquent d'avoir à limiter la quantité de matériaux qu'il doivent gérer. En ce qui concerne les stratégies utilisées pour fournir l'accès :

- il sera possible d'adopter une méthode d'accès minimale, en stockant de manière sûre les données avec une bonne documentation sur les moyens d'accès originaux, en espérant que dans un stade à venir on pourra utiliser ces informations pour imaginer un moyen d'accès ;
- il sera possible d'identifier les matériaux prioritaires qui pourraient être transférés dans un format qui fournisse au moins un niveau d'accès, tout en stockant l'original pour un travail de préservation ultérieur offrant une accessibilité plus large ;
- en insistant pour accepter des matériaux seulement dans quelques formats très bien normalisés et largement employés, on peut considérablement réduire le volume de travail et les outils spéciaux nécessaires pour fournir l'accès. Cela rend même possible de procéder au passage de certains matériaux de façon satisfaisante avec des outils de qualité grand public.

## ETUDES DE CAS

### 17.21 Des stratégies possibles pour différent types de données (à débattre)

- Ensembles de données : codage normalisé ; métadonnées décrivant la structure ; peuvent faire l'objet d'un passage, mais comptent sur le fait que les données restent lisibles pendant une longue période avant une action ultérieure.
- Bases de données : saisir les données et la structure qui les documentent par l'extraction des données ; saisir le logiciel utilisé pour interroger les données ; saisir l'interface et les vues instantanées des résultats de l'interrogation ; passage des données à une structure de base de données et à une interface d'utilisation nouvelles.
- Fichiers d'images et fichiers audio : utilisation des normes, avec une attention particulière envers les facteurs qui risquent de causer des complications comme la compression ; passage à un nouveau format normalisé.
- Fichiers de texte : codage (c'est-à-dire passage à un codage normalisé et à une structure normalisée XML) ; impression possible ; passage à un nouveau format.
- Dossier HTML : trier par types de format et procéder au passage au fur et à mesure que la norme de navigation change.
- Logiciels et matériaux logiciels : préservation technique ; émulation ; réagencement technique.
- Courriels : extraction des données et structuration normalisée ; passage quand c'est nécessaire.
- Enregistrements de bureau : visualiseurs ; extraction des données et normalisation du format.

## REFERENCES - pour en savoir plus

### Références internes

*Les éléments essentiels*, voir aussi Décider de ce que l'on conserve : chapitre 12.

*Le rapport entre les données et le logiciel*, Voir aussi Comprendre la préservation numérique : chapitre 8.

### Références externes

Le site *Web Preserving Access to Digital Preservation (PADI)* fournit un ensemble de références exhaustif et actualisé pour étudier les stratégies destinées à préserver l'accessibilité. Il est disponible à : <http://www.nla.gov.au/padi/>

## Chapitre 18. Quelques points de départ

### INTRODUCTION

#### 18.1 Les objectifs

Il est difficile d'établir des programmes de préservation exhaustifs en partant de zéro, même s'ils ont des moyens importants. C'est un véritable défi pour ceux qui disposent de peu de ressources. Ce chapitre vise à suggérer quelques points de départ, à la fois sur le plan global et pour répondre à un certain nombre de scénarios hypothétiques possibles. Ces informations sont proposées pour stimuler la réflexion et la discussion, car chaque situation de programme est différente et exige des réponses spécifiques et souvent imaginatives.

### LES PRINCIPALES QUESTIONS DE GESTION

#### 18.2 Quelques démarches pour commencer

Les démarches suivantes aideront peut-être à établir un programme de préservation :

1. déterminer le type de matériaux dont vous êtes responsable, ou celui que vous souhaitez préserver ;
2. vous lier avec d'autres qui ont des intérêts ou des responsabilités similaires pour voir si une méthode coopérative est possible ;
3. vous lier avec d'autres qui ont acquis une expérience dans la préservation ou du moins dans la gestion du type de matériaux qui vous intéresse, et chercher leurs conseils et orientations ;
4. essayer d'établir qui crée le matériau qui vous intéresse ; qui publie ou distribue ou détient le matériau, et quel intérêt et quelle compétence ces personnes peuvent avoir en le préservant pour au moins une période définie ;
5. essayer d'établir qui sont les utilisateurs potentiels et actuels du matériau, et comment ils souhaitent utiliser le matériau ;
6. il peut être trop difficile à ce stade d'identifier les caractéristiques essentielles à garder, mais il est important d'essayer de déterminer le niveau de fonctionnalité que vous voulez conserver, par exemple si les utilisateurs ont besoin de dialoguer avec le matériau et de le modifier, ou simplement de le visualiser sous la forme "lecture uniquement" ;
7. il semble qu'il y ait deux modèles largement utilisés pour les premières mesures pratiques :
  - commencer modestement par une petite quantité de matériau, en se bornant par exemple à un matériau simple et "clair", avec l'objectif de fournir le meilleur niveau de préservation possible dans le cadre de ces contraintes et au fur et à

mesure qu'on apprend, et en s'engageant à construire sur cette base une politique, des objectifs, des compétences et une infrastructure ;

- chercher à concevoir l'ensemble du programme et la manière dont les défis seront relevés avant de commencer ;

les deux méthodes ont des inconvénients et des avantages, mais ces directives recommandent la première méthode pour les personnes qui ont des ressources limitées et qui doivent répondre à des besoins urgents. Certes cette méthode peut difficilement prétendre qu'elle offre une préservation exhaustive et fiable, mais elle offre une préservation qui peut développer l'exhaustivité et la fiabilité avec le temps ;

8. développer au moins des politiques de base qui guideront les premiers engagements que vous prendrez ;
9. identifier les menaces flagrantes qui réclament une attention immédiate pour éviter que le matériau précieux soit perdu. (Il faut parfois accepter que certains matériaux soient perdus, et se consacrer à sauver au moins une partie des matériaux les plus importants) ;
10. identifier toutes les mesures immédiates que vous aurez à adopter pour mieux répondre à ces menaces, comme acquérir une connaissance plus approfondie du matériau que vous aurez à traiter, ou établir des contacts avec les créateurs ;
11. identifier les ressources - les gens, les compétences, les capitaux, l'équipement, le temps - qui peuvent être investies dans la réponse à fournir aux menaces ;
12. identifier les actions que vous pouvez entreprendre, en particulier les mesures simples qui peuvent être rapidement mises en oeuvre, soit pour vous faire gagner du temps, soit, si vous avez beaucoup de ressources et de soutien, vous permettre de vous embarquer dans un programme de préservation plus ambitieux. (Quelques exemples de menaces imminentes et d'actions à entreprendre figurent ci-après dans ce chapitre) ;
13. établir la liste des droits et des autorisations dont vous aurez besoin pour entreprendre cette action ;
14. limiter les autorisations, soit en clarifiant les droits existants, soit en demandant aux propriétaires des droits l'autorisation de faire ce que vous êtes sur le point de faire ;
15. planifier et entreprendre l'action que vous avez retenue et l'évaluer à chaque étape ;
16. discuter avec les créateurs de la manière dont ils créent leurs matériaux, en leur proposant des conseils et des informations sur les pratiques qu'ils pourraient suivre pour faciliter la préservation ;
17. passer en revue ce que vous avez fait et décider si c'est à poursuivre, à élargir ou à abandonner ; si un développement ultérieur ou une modification importante sont nécessaires.

## EXEMPLES DE SCENARIOS

### 18.3 Scénario 1

Pour les publications Web, il est fort probable qu'elles seront modifiées ou enlevées de la Toile sans avoir été sauvegardées.

*Actions possibles pour y répondre :*

- Développer les critères de sélection du matériau qui mérite d'être sauvegardé.
- Joindre les éditeurs et discuter de leurs projets et des vôtres.
- S'il semble que le matériau risque d'être perdu, et que vous pouvez obtenir l'autorisation, copiez les fichiers adéquats sur un ordinateur local, vérifiez la qualité du transfert, documentez ce que vous avez copié et la manière dont vous l'avez copié, et commencez à vous occuper du train de données, en faisant des copies de sauvegarde et en les stockant en lieu sûr hors site.
- Sur la base des engagements pris avec l'éditeur, décider s'il convient de rendre le matériau disponible à l'accès public. Si oui, vous aurez besoin d'une interface qui permette aux utilisateurs de le trouver et de comprendre ce qu'ils voient.
- C'est un dispositif à très court terme qui finira par avoir besoin du soutien de systèmes pour explorer les matériaux, gérer les négociations avec les éditeurs, saisir et télécharger une gamme de matériaux, enregistrer et gérer les métadonnées adéquates, gérer les conditions d'accès, surveiller l'entretien permanent des trains de données, identifier les caractéristiques essentielles du matériau qui doivent être conservées lorsque les technologies d'accès actuelles ne fonctionnent pas, et trouver des moyens de représenter le matériau qui peut inclure des multimédias complexes et des objets exécutables. Les systèmes et les dispositifs pour faire tout cela réclament des investissements très importants en temps et en argent, et leur mise en place peut prendre des années de développement et de travail d'approvisionnement.
- Il y a un certain nombre de programmes de préservation dans le monde qui développent ces aptitudes : il faut consulter leurs spécifications techniques, leurs politiques de préservation et leurs manuels de procédures.

### 18.4 Scénario 2

En ce qui concerne les enregistrements dans un système de conservation d'enregistrement, il peut se produire un changement imminent, par exemple un changement de gouvernement, qui entraînera un transfert à grande échelle ou la mise au rebut des enregistrements ou un remplacement imminent des systèmes.

*Les actions possibles pour y répondre :*

- S'il y a du temps, engager une étude d'évaluation pour déterminer comment fonctionne le système de conservation d'enregistrement, quelles activités de gestion sont enregistrées, quels types d'enregistrement sont importants et si le système de conservation d'enregistrement autorise de faire une présomption d'authenticité. Les enregistrements seront condamnés selon le calendrier de mise au rebut préparé à partir du processus d'évaluation, et les enregistrements sélectionnés pour la préservation seront transférés, avec leurs métadonnées, dans un dépôt d'archivage, où ils seront vérifiés, puis convertis dans un format normalisé que l'on aura choisi

pour sa capacité à contenir leurs caractéristiques essentielles, et enfin placés en stockage avec les sauvegardes appropriées.

- Si le temps manque pour opérer l'évaluation nécessaire avant le transfert, il faudra déterminer, par une évaluation rapide, s'il y a une documentation suffisante pour permettre un traitement ultérieur, une fois les matériaux retirés de leur contexte de travail. Les enregistrements et tous les outils qui permettent d'y accéder - éventuellement le système de conservation d'enregistrement tout entier - peuvent être transférés dans un dépôt d'archivage.
- Après le transfert, établir des dispositifs pour un stockage sûr des enregistrements, de leur documentation, et de tous les outils d'accès jusqu'à ce qu'on les évalue. De nombreux systèmes de conservation d'enregistrement étant liés à des plate-formes d'exploitation spécifiques, il faudra peut être conserver l'équipement et le logiciel de soutien originaux.
- Il est vraisemblable que de tels matériaux réclameront de façon urgente une estimation et la mise en place de stratégies adéquates pour conserver l'authenticité et l'accessibilité, telle qu'une extraction des données et le passage à un format normalisé, avant que le système de conservation d'enregistrement ne soit plus viable.

### 18.5 Scénario 3

Pour les matériaux audiovisuels produits commercialement, il peut y avoir un risque que les petits producteurs fassent faillite :

*Action possible pour y répondre :*

- Elaborer des critères pour sélectionner ce qui vaut la peine d'être conservé.
- Identifier les producteurs ou les distributeurs, entrer en contact avec eux et chercher à se procurer des copies des fichiers source de la meilleure qualité dépourvus de dispositifs anticopie.
- Si le matériau est dans un format non normalisé, demander des copies dans un format normalisé.
- Transférer le matériau qui a été sélectionné pour la préservation au centre d'archivage, et le copier soit sur un support plus stable, soit sur un système de stockage informatique bien géré avec les sauvegardes adéquates.
- Contrôler la qualité du transfert des données.
- Enregistrer la documentation sur le matériau, y compris les informations de droits susceptibles d'être nécessaires dans la gestion des droits de propriété intellectuelle, qui sont souvent compliqués pour ce type de matériau.
- Etant donné que les matériaux audiovisuels numériques comptent de très grandes quantités de données, il est peu probable qu'un centre d'archives non spécialisé aura les installations requises pour stocker de nombreux éléments avant de rencontrer des problèmes de capacité. Il est préférable de les stocker sur un support autonome stable comme un CD.
- Des programmes plus ambitieux doivent développer des dispositifs de dépôt et de droits avec les producteurs ou les distributeurs, et mettre au point des systèmes perfectionnés pour manipuler et stocker de grandes quantités de données, de métadonnées, et éventuellement un équipement et

un logiciel ésotériques de lecture s'il n'a pas été possible de convertir le matériau dans le système de lecture normalisé du centre d'archivage sans perte inacceptable.

#### **18.6 Scénario 4**

Pour les matériaux publiés sur des supports magnétiques à court terme comme les disquettes souples ou les bandes, il existe une forte probabilité de détérioration des médias.

*Action possible pour y répondre :*

- Chercher à déterminer quels matériaux méritent d'être préservés.
- Copier les matériaux des disquettes ou des bandes soit sur des supports plus stables comme les CD, soit sur d'autres supports instables comme des bandes qui sont gérées activement par des systèmes d'entretien de données.
- Vérifier la qualité de la copie.
- Enregistrer les métadonnées ayant trait aux matériaux et au transfert.
- Si possible, utiliser le transfert comme une occasion de documenter les dépendances logicielles du matériau.
- Planifier la manière de traiter le problème de dépendance de logiciel.

#### **18.7 Scénario 5.**

Pour les matériaux "hérités", il peut y avoir une perte imminente de l'équipement et du logiciel nécessaires pour l'accès.

*Action possible pour y répondre :*

- Chercher à déterminer quels matériaux méritent d'être préservés.
- Si possible, prendre des dispositions pour que le conservateur actuel transfère les matériaux sur un support et un format de fichier qui puissent être traités par l'équipement et le logiciel que vous avez, à condition que cela se fasse sans perte importante.
- Si nécessaire, prendre des dispositions pour que les matériaux dans leur état original soient transférés à votre garde avec l'équipement et le logiciel actuellement utilisés pour leur accès.
- Chercher une manière d'entretenir l'équipement aussi longtemps que possible, ou copier immédiatement les matériaux sur un support et/ou un format différents.
- Documenter le matériau, sa provenance et toutes les modifications.
- Stocker sûrement les matériaux avec les sauvegardes adéquates.
- Planifier la manière dont vous allez agir avec les dépendances logicielles des matériaux, notamment s'il n'a pas été possible de les convertir dans un format de fichier que vous continuerez d'utiliser.

- Pour les matériaux qui ont déjà perdu le matériel et le logiciel nécessaires à l'accès, il sera probablement nécessaire de trouver quelqu'un possédant le même équipement en état de fonctionnement qui vous permettra de l'utiliser. Cela peut nécessiter l'utilisation de services ésotériques de récupération de données ou l'achat d'un logiciel spécialisé pour la récupération des données.

## **18.8 Scénario 6.**

Pour les ensembles de données complexes, il peut y avoir une perte imminente du personnel qui sait comment les données sont codées et comment fonctionne l'ensemble de données.

*Action possible pour y répondre :*

- Déterminer si l'ensemble de données mérite d'être gardé.
- Avec l'aide du personnel existant qui connaît bien l'ensemble de données, documenter cet ensemble de manière approfondie de sorte qu'un autre personnel ou des fournisseurs puissent reprendre sa gestion.
- S'assurer que l'ensemble de données est copié et convenablement sauvegardé dans un système de stockage sûr.
- Planifier la manière dont vous continuerez à fournir l'accès une fois que l'environnement d'exploitation actuel sera dépassé.

## **PROGRAMMES DE PRESERVATION MINIMAUX**

Cette partie propose quelques mesures possibles pour établir un programme de préservation minimal. Elles sont applicables dans les cas où une action est nécessaire, sans que l'Organisation soit en mesure de s'engager dans quelque chose de plus ambitieux. Figurent aussi quelques scénarios.

### ***Comprendre vos responsabilités de préservation, vos besoins et vos ressources***

- Y a-t-il des matériaux numériques que vous devez préserver ? Y a-t-il quelqu'un susceptible de les préserver ? Quelles sont les autorisations nécessaires ? Contre quels risques et quelles menaces doit-on se prémunir ? Déterminer les ressources que vous pouvez consacrer à cette tâche.

### ***Influencer la tâche de préservation***

- Décider au moins des formats qui seront acceptés. Si possible, négocier avec les producteurs pour qu'ils utilisent largement les normes admises et qu'ils fournissent la documentation adéquate.

### ***Protéger les données***

- Stocker les supports dans des conditions appropriées.
- Copier les données sur des supports plus stables et faire des copies de sauvegarde, en utilisant des supports de bonne qualité.

- Stocker les données de manière sûre et, si possible, choisir un stockage hors site pour les sauvegardes.
- Rechercher régulièrement les erreurs dans les données.
- Etablir un régime de rafraîchissement des données adapté à la vie des supports.

#### ***Agir en ce qui concerne les moyens de fournir l'accès***

- Enregistrer les informations qui seront nécessaires pour fournir un accès à court terme : l'identité du matériau, les besoins d'accès, les mots de passe, etc.
- Conserver l'équipement et le logiciel d'accès nécessaires, entretenir le matériel et protéger le logiciel dans le cadre des dispositifs de licence.
- Une autre action sera nécessaire pour conserver l'accessibilité lorsqu'on ne pourra plus maintenir le matériel en état de fonctionnement.
- Si une action supplémentaire n'est pas réalisable pour les ressources du programme minimal, prévoir de transmettre le matériau à un autre gardien qui en prendra la responsabilité.
- Ou, autre solution, trouver des moyens de reproduire de façon appropriée le matériau sous une forme non numérique qui soit stable (comme l'impression). Cette solution risque d'être peu satisfaisante pour tout ce qui n'est pas texte ou images fixes.

#### **Un programme minimal - Scénario 1**

Une petite maison de production conserve des archives de ses fichiers à chaque travail d'édition. Ces archives servent à fournir des fichiers, source de contenu et de référence en vue d'une réutilisation du contenu pour ses publications. La société a besoin d'accéder au matériau archivé à son propre usage pendant au moins cinq ans. Un petit nombre de ses publications tombent sous le coup du dépôt légal et doivent être déposés sous forme numérique auprès d'une institution de collecte pour la préservation à long terme. La société n'assume plus de responsabilité de préservation pour son matériau.

Les archives de la société sont stockées sur CD-ROM dans une zone sûre au sein même de l'entreprise, avec une copie supplémentaire de chaque CD-ROM stockée hors site. Les premiers fichiers d'archives ont été stockés sur des supports magnétiques et magnéto-optiques. Le matériau figurant sur ces disques a été transféré sur CD-ROM, car le support magnétique risque de ne pas être assez stable pendant toute la période requise et les lecteurs spécialisés nécessaires pour lire les disques sont progressivement retirés de la vente.

Les archives comprennent à la fois les fichiers de préproduction et de production dans des formats logiciels d'édition exclusifs et spécialisés. La société compte sur une compatibilité en amont ou sur des versions modernisées de logiciels pour conserver l'accès à ses anciens fichiers. Pour les fichiers d'image critique et de référence de conception, des versions de références sont également créées dans des formats largement restituables comme le TIFF ou le PDF pour un accès élargi. Ces versions sont prévues comme des références pour recréer les fichiers de travail *de novo* dans le futur, si les versions ultérieures du logiciel n'étaient plus en mesure de restituer les anciennes versions.

## **Un programme minimal - Scénario 2**

Une institution de recherche parraine un programme de patrimoine d'une communauté indigène. Les membres de la communauté créent une collection d'enregistrements sur le folklore, l'art, la généalogie et d'entretiens, qu'ils rendent disponible sur l'Internet. Le matériau est hébergé par l'institution de recherche et est de valeur durable en tant qu'enregistrement de la culture et des traditions de la communauté indigène. Des droits ont été négociés avec la communauté pour permettre au matériau d'être préservé indéfiniment.

Le site utilise des fichiers de médias et de balisage pour le Web actuels et normalisés, à la demande de l'institution d'accueil, qui assume la responsabilité initiale de préservation pour le site, et conserve une petite collection d'outils logiciels qui peuvent être utilisés pour rendre de manière correcte le contenu qu'elle abrite.

L'institution entretient régulièrement les copies de sauvegarde de tous les matériaux hébergés sur ses serveurs, y compris le stockage hors site d'au moins une copie de sauvegarde. Les sites individuels peuvent aussi être régulièrement archivés sur des CD-ROM, notamment chaque fois qu'on les actualise. Les mesures de sécurité normalisées des techniques de l'information sont en place pour protéger le contenu en ligne.

L'institution dispose de financements pour poursuivre le programme pendant deux nouvelles années et est en train de négocier le transfert des matériaux à une autre institution pour une préservation continue.

## **Un programme postminimal - Scénario 3**

Une institution d'enseignement supérieur et de recherche entretient un programme d'archivage des mémoires, des ensembles de données de recherche, des analyses et des modèles. Elle prévoit de préserver et de fournir l'accès à ces matériaux indéfiniment.

L'institution limite la gamme des formats dans lesquels seront acceptés les matériaux déposés, afin de réduire le champ des voies de la préservation qu'il faut concevoir sur le long terme. Les formats acceptés sont conformes aux normes ouvertes, très répandues, dont on attend qu'elles soient accessibles pendant une période plus longue. L'institution conserve un logiciel capable de restituer ces formats aujourd'hui. On demande aux déposants de fournir des métadonnées détaillées.

L'institution fournit un stockage sûr, à la fois sur batterie de disque et sur bande, et tous deux ont en place un régime de sauvegarde et de mesures anticatastrophe, y compris des copies multiples dans des endroits séparés et sur des supports différents.

On est en train de concevoir les plans de modernisation du système et de l'équipement, et de chercher des méthodes pour traduire ou interpréter dans l'avenir les formats déposés. L'environnement technologique est également surveillé afin d'y détecter les signes d'obsolescence imminente des technologies utilisées.

**QUATRIEME PARTIE**  
**INFORMATIONS COMPLEMENTAIRES**

## **19 Glossaire**

### **19.1 Les objectifs**

Ce glossaire sélectif explique le sens dans lequel certains termes sont employés dans ces directives.

### **19.2 Les termes**

**Accessibilité** : La capacité d'accéder au sens essentiel, authentique ou à la finalité d'un objet numérique.

**ASCII** : American Standard Code for Information Interchange ("code normalisé américain pour les échanges de données"). Code normalisé utilisé internationalement pour représenter toutes les majuscules et minuscules de l'alphabet latin, les nombres, la ponctuation, etc.

**Authenticité** : Qualité de véridicité de certains matériaux numériques, qui se présentent comme étant ce qu'ils sont censés être, soit en tant qu'objet original, soit en tant que copie fiable obtenue par des processus parfaitement documentés à partir de l'original.

**Bit** : (Binary Digit) (chiffre binaire). La plus petite unité de quantité d'information informatisée étant un seul chiffre (1 ou 0).

**Journal intime sur le Web** : (Weblog). Un journal de messages envoyés à un site Web, souvent par le propriétaire du site, mais souvent aussi par des correspondants invités.

**Camirette** : Sites Web qui diffusent des images depuis une caméra vidéo reliée à un ordinateur, soit sous la forme d'une succession d'images fixes, soit sous la forme d'une vidéo transmise en continu.

**Certification** : Processus d'évaluation du degré de conformité d'un programme de préservation avec un ensemble agréé de normes ou de pratiques minimales.

**Compression** : Réduction de la quantité de données nécessaires pour stocker, transmettre et représenter un objet numérique

**Navigateur** : Logiciel qui fournit l'accès aux pages de la Toile.

**Octet** : Ensemble de huit bits représentant un seul caractère dans le code informatique.

**Objets Conceptuels** : Objets numériques avec lesquels les humains communiquent sous une forme intelligible à l'homme.

**Protection des données** : Processus qui vise à protéger les données de niveau binaire des objets numériques contre les changements non autorisés ou les pertes.

**Patrimoine numérique** : Les matériaux numériques qui présentent assez de valeur pour être conservés en vue d'un accès et d'une utilisation futurs.

**Préservation numérique** : Les processus qui visent à conserver l'accessibilité des objets numériques dans le temps.

**Dispositifs répartis** : Des dispositifs pour la préservation numérique qui répartissent la responsabilité entre plusieurs partenaires.

**Télécharger** : Processus qui consiste à copier les données d'un ordinateur éloigné sur un stockage informatique local.

**DTD** : Document Type Définition. Fonction du langage SGML permettant de définir le contenu des documents.

**Cryptage** : Processus qui consiste à coder les données dans un code secret de sorte que seuls les utilisateurs autorisés peuvent reconvertir les données dans leur codage original pour l'affichage.

**Prétraitement électronique** : Textes numériques de travaux de recherche rendus accessibles par l'Internet.

**Éléments essentiels** : Les éléments, les caractéristiques et les attributs d'un objet numérique donné, qui doivent être préservés afin de représenter son sens ou son propos essentiel. Également appelés *propriétés significatives* par certains chercheurs.

**HTML** : (Hyper Text Markup Language). Le codage utilisé pour créer des pages sur la Toile, qui comprend des balises pour le formatage des textes, l'insertion d'objets et les liens hypertextuels.

**Identité** des objets de données : L'état distinctif de l'objet, qui le différencie d'autres objets numériques, y compris d'autres versions ou copies du même contenu.

**Ingestion** : Processus qui consiste à mettre les objets numériques et la documentation qui leur est associée dans un stockage sûr.

**Intégrité** des objets de données : L'état complet, intact et dépourvu de modifications non autorisées et non documentées.

**L'Internet** : Le plus grand ensemble mondial de réseaux interconnectés (ou le Réseau des réseaux), qui utilise tous les protocoles TCP/IP (Transmission Control Protocol/Internet Protocol).

**Objets logiques** : Les objets numériques comme le codage informatique, les objets conceptuels sous-jacents.

**Moyens d'accès**. Outils (en particulier, les combinaisons particulières de logiciels et de matériel) nécessaires pour fournir l'accès aux objets numériques et les présenter sous une forme lisible pour l'homme.

**Métadonnées** : Données sur les données, généralement sous une forme très structurée et souvent codées pour le traitement informatique et l'interrogation.

**Publications en ligne** : Documents numériques auxquels les utilisateurs peuvent accéder par le moyen d'un réseau informatique comme l'Internet.

**Normes ouvertes** : Spécifications qui sont définies au moyen d'un processus disponible au public et qui sont disponibles au public pour la consultation et l'utilisation.

**Système d'exploitation** : Logiciel qui contrôle la manière dont fonctionne un ordinateur.

**Objets physiques** : Objets numériques en tant que phénomènes physiques qui enregistrent le codage logique, tels que l'état de polarité des supports magnétiques, ou l'état de réflectivité des supports optiques.

**Modules d'extension** : Eléments de logiciel (parfois de matériel) qui ajoutent des fonctions à un programme logiciel plus étendu comme les programmes qui présentent des types de fichier spécifiques.

**Portage** : Processus qui consiste à transférer un élément de logiciel appartenant à un système informatique dans un autre système informatique.

**Producteurs** : Agents responsables de la conception, de la création et de la distribution des matériaux numériques.

**Métadonnées de préservation** : Métadonnées destinées à soutenir la gestion de la préservation des matériaux numériques, en documentant leur identité, leurs caractéristiques techniques, leurs moyens d'accès, leur responsabilité, leur histoire, leur contexte, leur objectif historique et leur objectif de préservation.

**Programme de préservation** : L'ensemble des dispositifs, et les personnes qui en sont responsables, mis en place pour gérer les matériaux numériques en vue d'une accessibilité permanente.

**Logiciels du domaine public** : Programmes logiciels qui ne sont pas soumis aux contraintes du copyright.

**Rafraîchissement** : Processus qui consiste à copier les données d'un support sur un autre, sans changer la manière dont les données sont codées, afin d'éviter une perte des données due à la détérioration du support ou à son remplacement.

**Métadonnées de découverte de ressource** : Métadonnées destinées à rendre visibles l'existence et la description des matériaux numériques à ceux qui souhaiteraient y accéder.

**Droits** : Autorisations juridiquement contraignantes associées aux matériaux numériques, comme le copyright, le respect de la vie privée, la confidentialité, et les restrictions en matière de sécurité nationale et internationale.

**Gestion du risque** : Processus qui consiste à identifier et à évaluer les risques présentés par des menaces, et s'il le faut, à prendre des mesures pour rabaisser le niveau de risque à un niveau acceptable.

**Prestataires de services** : Des organismes ou des personnes avec lesquels on a passé contrat pour remplir certaines ou toutes les fonctions d'un programme de préservation, sous la responsabilité prédominante du programme.

**SGML** : (Standard Generalized Markup Language). Une norme qui définit un langage de balisage pour les documents. Le SGML décrit la manière de définir (dans un DTD [Document Type Definition]) la structure sous-jacente d'un type de documents donné, sans définir comment ils seront présentés. Le HTML et le XML sont fondés sur le SGML.

**Normes** : Spécifications ou pratiques convenues pour atteindre des objectifs donnés. Certaines normes sont formellement préparées, agréées, avalisées et publiées par des organismes chargés d'établir les normes, tandis que d'autres deviennent des normes *de facto* par une adoption et une utilisation courantes. Certaines normes, comme de nombreux formats de fichier, sont développées et brevetées par les détenteurs de la propriété intellectuelle qui peuvent ou non rendre leurs spécifications publiques.

**Vérification** : Processus qui consiste à vérifier qu'un objet numérique dans un format de fichier donné est complet et qu'il est conforme à la spécification du format.

**World Wide Web** : (La Toile mondiale) L'ensemble des ressources et des serveurs accessibles par l'Internet utilisant le protocole HTTP, qui n'est qu'un seul des nombreux moyens permettant d'accéder à l'information par l'Internet. (Le courriel en est un autre.)

**XML** : (Extensible Markup Language). Version simplifiée du langage SGML qui est en passe de devenir une norme largement employée pour décrire les structures de documents normalisées afin qu'ils puissent être compris par la plupart des systèmes informatiques.

## **20. Bibliographie**

### **20.1 Les objectifs**

Cette bibliographie propose des sources d'informations supplémentaires qui entrent bien davantage dans le détail qu'il n'était possible pour ces Directives. Les lecteurs doivent être avertis que les analyses et les méthodes de la préservation numérique ne sont ni convenues ni établies universellement, aussi ces lectures peuvent-elles présenter des points de vue différents sur certaines questions.

### **20.2 Le contenu**

Il s'agit d'une bibliographie très sélective. Les informations disponibles sur la préservation numérique sont vastes ; le nombre des informations dans les domaines associés qui s'intéressent aux programmes de préservation, comme les métadonnées de découverte de ressource, les métadonnées de gestion de droits est considérable. Toutefois, quelques "textes standard" peuvent former le coeur de cette bibliographie. La majeure partie des références qui figurent ci-dessous sont extraites de la passerelle PADI (référéncée elle-même sous *Des sources de connaissances actuelles*).

Pour que la longueur de la liste des textes à lire reste gérable :

- la plupart des références figurant déjà dans les chapitres individuels n'ont pas été reprises ici. Elles doivent toutefois être considérées comme une partie importante (voire la plus importante) de la liste ;
- les communications de colloques ont été omises pour la plupart, car elles sont fort nombreuses et ont généralement une orientation plus théorique. Une recherche sur la passerelle PADI en utilisant le terme "conference paper" fournit l'accès à de nombreuses ressources utiles ;
- en général, on a privilégié les références à des sites de projet et d'organisation au détriment des textes multiples (et souvent répétitifs) qui les décrivent. Dans la plupart des cas, on peut trouver des textes qui s'y rapportent, via les sites eux-mêmes, ainsi que d'autres informations techniques.

### **20.3 Des sources de connaissances actuelles**

Les sites d'informations suivants sont de bonnes sources d'informations actualisées et de débat autour des problèmes de préservation numérique :

## **Sites Web**

PADI (Preserving Access to Digital Information). Bibliothèque nationale d'Australie. (Régulièrement mis à jour.)

PADI est une passerelle thématique internationale consacrée au sujet de la préservation numérique. En partenariat avec la Digital Preservation Coalition, elle publie également un résumé trimestriel des nouveaux développements importants. <http://www.nla.gov.au/padi/>

ERPANET: Electronic Resource Preservation and Access NETWORK. (Régulièrement mis à jour.) ERPANET, fondé par la Commission européenne, vise à établir un consortium international pour fournir un centre d'information virtuel et une base de connaissances sur l'état actuel de la technique dans la préservation numérique. <http://www.erpanet.org/>

## **Revue spécialisées (et série de monographies régulièrement publiées)**

*Parutions du CLIR*. Council on Library and Information Resources. (Régulièrement mises à jour.) <http://www.clir.org/pub/issues/issues.html>

*Rapports du CLIR*. Council on Library and Information Resources. (Paraissent régulièrement.) <http://www.clir.org/pubs/reports/reports.html>

*Revue D-Lib*. (Régulièrement mise à jour.)

Un journal en ligne de recherche numérique. <http://www.dlib.org/>

Digital Document Quarterly. Henry Gladney. (Régulièrement mis à jour.)

Ce journal publié par un particulier, traite des problèmes de préservation numérique. <http://home.pacbell.net.hgladney/ddq.htm>

Publications du European DigCULT Forum. (Paraissent régulièrement.) <http://www.digicult.info/pages/publications.php>

*RLG DigiNews*. Research Libraries Group.

Ce bulletin bimensuel fournit des informations sur les initiatives numériques qui présentent une composante ou une analyse raisonnée de préservation, sur des projets de conversion d'images et d'archivage numérique, et publie des annonces. Archives disponibles. <http://www.rig.org/preserv/diginews/>

## **Listes de discussion : l'essentiel**

DIGITAL-PRESERVATION. JISC

Véhicule des annonces et des informations sur des activités concernant particulièrement la préservation et la gestion des matériaux numériques au Royaume-Uni. <http://www.jicsmail.ac.uk/lists/digital-preservation.html>

OAIS Implementers Discussion List ([oais-implementers@lists2.rlg.org](mailto:oais-implementers@lists2.rlg.org))

Une liste de discussion destinée aux particuliers et aux institutions travaillant activement avec l'Open Archival Information Systems (OAIS) Reference Model pour essayer d'élaborer, édifier et gérer des archives ou des gisements de données numériques. <http://www.rig.org/longterm/oais.html>

#### padi/forum-I

padiforum-I est une liste de discussion régulée pour l'échange de nouvelles et d'idées relatifs aux problèmes de préservation numérique. <http://www.nla.gov.au/padi/forum>

#### WEB-ARCHIVE

Comité réseau des universités

Centré sur l'archivage en ligne du contenu, d'un point de vue technique, juridique et organisationnel. <http://listes.cru.fr/www/info/web-archive>

#### Listes de discussion : non essentielles mais utiles

##### ERECS-L

Une liste régulée pour les archivistes et autres professionnels de l'information, qui fournit un forum de discussion d'idées, de techniques et de problèmes relatifs à la gestion et à la préservation des enregistrements électroniques. <http://listserv.albany.edu:8080/archives/erecs-I.html>

##### Preservation Administration Discussion Group (PADG-L)

Cette liste recouvre la préservation à la fois des matériaux traditionnels et des matériaux numériques. Un service d'archives est consultable. <http://palimpsest.stanford.edu/byform/mailling-lists/padg>

##### ShelfLife

Un résumé hebdomadaire d'articles destiné aux cadres pour les professionnels de l'information du monde entier publié par le Research Libraries Group (RLG) en collaboration avec NewScan, Inc. <http://www.rlg.org/shelflife/index.html>

##### IASA list

Liste de discussion de l'International Association of Sound and AudioVisual Archives. <http://www.rlg.org/shelflife/index.html>

#### 20.4 D'intérêt général

CAMiLEON: Creative Archiving at Michigan and Leeds: Emulating the Old on the New. (Régulièrement mis à jour.)

Examine les problèmes relatifs à la mise en oeuvre de l'émulation de la technologie en tant que stratégie de préservation numérique, et souhaite développer des outils, des directives et l'établissement de coûts pour l'émulation par rapport à d'autres stratégies de préservation numérique. <http://www.si.umich.edu/CAMILEON/>

Cedars: Curl Exemplars in Digital Archives Project. (Mis à jour jusqu'en 2002.)

Sous la direction globale du Consortium of University Research Libraries, le projet (avril 1998 - mars 2002) vise à traiter les problèmes stratégiques, méthodologiques et pratiques de la préservation numérique. Le site Web est relié aux documents du Cedars Guidance relatifs aux droits de propriété, aux métadonnées de préservation, à la gestion de collection, aux stratégies

techniques et au Cedars Distributed Digital Archiving Prototype System, ainsi qu'aux articles et aux communications de travail du projet. <http://www.leeds.ac.uk/cedars>

Changing Trains at Wigan: Digital Preservation and the Future of Scholarship. Seamus Ross. (Date de création : novembre 2000.)

S'intéresse à l'émergence des matériaux documentaires numériques à des fins universitaires et de testimoniales, et étudie les défis et les problèmes que posent leur préservation efficace dans une perspective d'étude de cas. <http://www.bl.uk/services/preservation/ocpaper.pdf>

A Continuing Access and Digital Preservation Strategy for the Joint Information Systems Committee (JISC) 2000-2005. Neil Beagrie. (Date de création : 1er décembre 2002)

Présente le rôle que le JISC assumera pour le compte des conseils d'investissements et des institutions en tant que partie d'un programme national de préservation numérique. [http://www.jisc.ac.uk/index.cfm?name=press\\_continuing](http://www.jisc.ac.uk/index.cfm?name=press_continuing)

Cyberculture, Cultural Asset Management, and Ethnohistory: Preserving the Process and Understanding the Past. Seamus Ross. (Date de création : juin 2001.)

Souligne l'importance de la préservation du contexte culturel dans lequel fonctionne l'Internet, en mettant en lumière huit défis à relever pour assurer un accès à long terme des matériaux dans le cyberspace, et compare les avantages des modèles d'archivage centralisés, décentralisés et répartis. [http://www.deflink.dk/upload/doc\\_filer/doc\\_alle/740\\_sross\\_cyberculture\\_rev2.doc](http://www.deflink.dk/upload/doc_filer/doc_alle/740_sross_cyberculture_rev2.doc)

Digital Division is Cultural Exclusion. But Is Digital Inclusion Cultural Inclusion? Karen Worcman in: *D-Lib Magazine*. (Date de création : mars 2002.)

Etudie "jusqu'où les technologies numériques et l'Internet peuvent être les instruments d'une intégration sociale et culturelle" et "comment l'utilisation de ces technologies peut être liée à la préservation de l'histoire d'un groupe culturel particulier". Note aussi les conséquences de la technologie numérique sur l'histoire et la mémoire collective des communautés et les défis à relever pour vaincre l'exclusion numérique des groupes économiquement défavorisés, pour créer et préserver l'histoire numérique, ainsi que des projets et des ressources durables. <http://www.dlib.org/dlib/march02/worcman/03worcman.html>

Digital Electronic Archiving: the State of the Art and the State of the Practice. B. C. Carroll, G. Hodge; Information International Associates Inc. (Date de création : 26 avril 1999.)

Etude réalisée pour fournir des informations sur l'état de la technique et des pratiques dans les politiques d'archivage électronique numérique, les modèles et les meilleures solutions. D'envergure internationale, elle inclut une variété de types de données applicables aux informations techniques et scientifiques y compris des données, du texte, des images, du son, de la vidéo et du multimédia, et une variété de types d'objets comme les journaux et les monographies électroniques, les images par satellite, les données de séquence biologiques et les brevets. Plusieurs projets "de pointe" sont retenus pour une analyse plus détaillée. [http://www.icsti.org/99ga/digarch99\\_TOCP.pdf](http://www.icsti.org/99ga/digarch99_TOCP.pdf)

Digital Preservation and Deep Infrastructure. Stewart Granger in : *D-Lib Magazine*. (Date de création : février 2002.) <http://www.dlib.org/dlib/february02/granger/02granger.html#>

European Commission on Preservation and Access (ECPA). (Régulièrement mis à jour.)

L'ECPA fait fonction de "plate-forme européenne pour la discussion et la coopération des organisations du patrimoine dans les domaines de la préservation et de l'accès". Le site Web renferme des informations sur les projets, les activités, les publications et les autres ressources liées à la préservation du patrimoine documentaire (y compris les matériaux numériques) en Europe. <http://www.knaw.nl/ecpa/about.html>

JISC Digital Preservation Focus. Joint Informations Systems Committee. (Régulièrement mis à jour.) <http://www.jisc.ac.uk/dner/preservation/>

Levels of Service for Digital Repositories. William LeFurgy in: *D-Lib Magazine*. (Date de création : mai 2002.)

William LeFurgy de l'US National Archives and Records Administration (NARA) souligne les conditions qui régissent la permanence des objets numériques, telles que l'architecture du système et la spécification des matériaux, et propose un modèle de niveaux de service futurs pour les gisements de données numériques. <http://www.dlib.org/dlib/may02/lefurgy/05lefurgy.html>

A Metadata Approach to Preservation of Digital Ressources: The University of North Texas Libraries' Experience. Cathy Nelson Hartman, Daniel Gelaw Alemneh, Samantha Kelly Hastings. (Date de création : août 2002.)

Cette communication discute des problèmes liés à la préservation numérique et montre le rôle que joue la préservation des métadonnées pour faciliter les activités de préservation en général. Elle décrit, notamment, les efforts engagés par les bibliothèques de l'UNT (University of North Texas) pour assurer l'accès et la préservation à long terme des diverses ressources de l'information numérique. [http://www.firstmonday.org/issues/issues7\\_8/alemneh/index.html](http://www.firstmonday.org/issues/issues7_8/alemneh/index.html)

Preserving Digital Information: Final Report and Recommendations. John Garrett, coprésident ; Task Force on Archiving of Digital Information; Donald Waters (président). (Date de création : 20 mai 1996.)

Résulte d'une décision de la Commission on Preservation and Access (CPA) et du Research Libraries Group (RLG) de charger un groupe spécial d'étudier et de recommander des moyens d'assurer "l'accès continu indéfiniment dans le futur d'enregistrement stockés sous une forme électronique numérique". Cet exercice qui a marqué un tournant décisif a provoqué une discussion dans le monde entier. <http://www.rlg.org/ArchTF/>

The state of the art and practice in digital preservation. Kyong-Ho Lee, Oliver Slattery, Richang Lu, Victor McCrary, Victor Tang. (Date de création : janvier 2002.)

Cette communication publiée dans le *Journal of Research of the National Institute of Standards and Technologies* (vol. 107, n° 1, p. 93-106) passe en revue les idées et les pratiques jusqu'à la fin 2001. Une dernière partie recommande le développement de normes de préservation fondées sur le XML, et souligne les problèmes cruciaux qui restent à résoudre. <http://nvl.nist.gov/pub/nistpubs/jres/107/1j71lee.pdf>

## **20.5 Une mobilisation pour la préservation**

Digital Preservation Coalition, JISC Digital Preservation Focus. (Dernière mise à jour : 14 février 2001.)

Fondée en 2001, la Digital Preservation Coalition (Royaume-Uni) vise à développer et à poursuivre un programme de préservation numérique anglais dans un contexte international. <http://www.dpconline.org/>

## 20.6 La préservation des matériaux publiés

Access to Web archives: le Nordic Web Archive Acces Project. Svein Arne Brygfjeld. (Date de création : 22 août 2002.)

Présenté à la 68e Conférence générale de l'IFLA, Glasgow, 2002. Décrit un système prototype pour les archives Web à grande échelle, tel qu'il a été développé par le Nordic Web Archive Access Project, une initiative des bibliothèques nationales du Danemark, de la Finlande, de l'Islande, de la Norvège et de la Suède. <http://www.ifla.org/IV/ifla68/papers/090-163e.pdf>

Egalement disponible en français à : <http://www.ifla.org/IV/ifla68/papers/090-163f.pdf>

Berkeley Digital Library SunSITE. University of California Berkeley Library and Sun Microsystems Inc. (Régulièrement mis à jour.)

Ce site crée des collections et des services numériques, tout en fournissant des informations et un soutien aux programmeurs de bibliothèque numérique dans le monde entier. Il comprend des liens avec les informations sur le copyright, les métadonnées, la préservation et les normes ; les projets de bibliothèque numérique ; les outils pour créer des bibliothèques numériques ; et une formation pour les bibliothécaires numériques. <http://sunsite.berkeley.edu/>

Collecting and Preserving the Web: Developing and Testing the NEDLIB Harvester. Juha Hakala, dans : *RLG DigiNews*. (Date de création : avril 2001.)

Fait part des résultats du NEDLIB Harvester Project for the Archiving of the Web Ressources. Rend compte de quelques problèmes essentiels concernant l'utilisation de cette forme de technologie pour saisir des matériaux sur la Toile. <http://www.rlg.org/preserv/diginews/diginews5-2.html#feature2>

Columbia University Libraries Policy for Preservation of Digital Resources. (Date de création : juillet 2000.)

Exposé d'une politique, notamment l'engagement dans la gestion du cycle de vie numérique. <http://www.columbia.edu/cu/lweb/services/preservation/dlpolicy.html>

DACHS: Digital Archive for Chinese Studies. Institut d'Etudes Chinoises, université de Heidelberg. (Régulièrement mis à jour.)

Opérationnel depuis août 2001, le DACHS "vise à identifier, archiver et rendre accessibles les ressources de l'Internet concernant les études chinoises, en insistant sur le discours social et politique, tel qu'il est reflété par les articulations de l'Internet chinois". Les ressources collectées comprennent les sites Web, les comités de discussion, les revues spécialisées, les bulletins d'information et les documents simples. La politique de collection du service d'archives, le déroulement des opérations et l'infrastructure technique font l'objet d'une présentation d'ensemble qui est disponible. <http://www.sino.uni-heidelberg.de/dachs/>

DELOS Network of Excellence (NoE) on Digital Libraries. (Régulièrement mis à jour.)

Créé en 2000 pour faciliter le développement d'un programme ouvert pour la recherche des bibliothèques numériques. Le groupe est un point de référence pour tous les projets du 5th Framework Programme fondés par l'Information Societies Technologies (IST) Programme. <http://www.ercim.org/delos/>

Digital Imaging and Preservation Policy Research(DIPPR). Department of Preservation and Conservation, Cornell University Library. (Dernière mise à jour : 22 mai 2002.)

Les membres du DIPPR sont issus du Department of Preservation and Conservation de la Bibliothèque de Cornell University et sont engagés dans la recherche, l'exécution, la publication et la formation, l'accent étant mis sur la préservation numérique et l'aspect prioritaire des résultats des projets de recherche permanents. Ses activités comprennent la recherche sur les aspects techniques de la numérisation d'images, la recherche de la préservation numérique à travers le Project Prism, et la publication de *RLG DigiNews*. <http://www.library.cornell.edu/iris/research/dippr.html>

Dspace: Durable, Digital, Depository. (Dernière mise à jour : 2002.)

Site Web installé à l'initiative de MIT Dspace avec Hewlett-Packard pour créer des archives électroniques de la contribution intellectuelle du MIT et d'autres partenaires fédérés basées sur la Toile. Les détails relatifs au recrutement du personnel, à la direction, à la planification, à l'architecture technique et à la fonctionnalité sont disponibles sur le site Web <http://dspace.org/index.html>

A First Experience in Archiving the FrenchWeb. Serge Abiteboul, Gregory Cobéna, Julien Masanès, Gerald Sedrati. (Date de création : septembre 2002.)

Décrit le travail préliminaire de la Bibliothèque nationale de France et de l'INRIA pour l'archivage de la Toile française soumise au dépôt légal. Débat également la définition du périmètre de la Toile française et les problèmes de gestion des versions. <ftp://ftp.inria.fr/INRIA/Projects/verso/gemo/GemoReport-229.pdf>

IFLANET : International Federation of Library Associations and Institutions. (Régulièrement mis à jour.)

Le site comporte une vaste gamme d'informations comprenant des ressources sur les collections et les services électroniques <http://www.rlg.org/preserv/diginews/diginews6-3.html#interview>

Internet Archive (Régulièrement mises à jour.)

Une initiative commerciale non lucrative qui collecte et stocke des matériaux publics en provenance de l'Internet comme la Toile mondiale, le Netnews et le logiciel téléchargeable donné par Alexa Internet. Les pages Web archivées sont accessibles en utilisant l'interface Wayback Machine. Fournit également l'accès à des films documentant la culture et la vie nord-américaine du XXe siècle, numérisés à partir des Prelinger Archives des films éphémères à San Francisco. <http://www.archive.org/>

Voir également : The Internet Archive, an Interview with Brewster Kahle, in: *RLG DigiNews*. (Date de création : 15 juin 2002.) <http://www.rlg.org/preserv/diginews/diginews6-3.html#interview>

JSTOR (Journal Storage): The Scholarly Journal Archive. (Régulièrement mis à jour.)

Vise à créer des archives fiables et exhaustives de travaux d'érudition importants parus dans des revues spécialisées. Les anciens numéros des revues spécialisées ont été convertis dans des formats

électroniques qui permettent des économies d'espace tout en améliorant l'accès au contenu de la revue. <http://www.istor.org/>

Kulturaw3. Bibliothèque nationale de Suède. (Régulièrement mis à jour.)

L'objectif de ce projet est de tester les méthodes de collecte, de préservation et de fourniture d'accès aux documents suédois en ligne. <http://kulturaw3.kb.se/html/kulturaw3.eng.html>

The Last Page of the Internet: the Importance of Preserving the Dynamic Aspects of the Internet. Niels Bruggen. (Date de création : juin 2001.)

Discute des complications qu'entraîne la préservation des caractéristiques dynamiques de l'Internet, selon les vues du spécialiste des médias Niels Bruggen. [http://www.deflink.dk/upload/doc\\_filer/doc\\_alla/1023\\_NBR.doc](http://www.deflink.dk/upload/doc_filer/doc_alla/1023_NBR.doc)

LOCKSS. Bibliothèques de l'Université de Stanford. (Régulièrement mis à jour.)

Le projet est en train de créer un logiciel d'"accès permanent" pour les bibliothèques. Le LOCKSS (Lots of Copies Keeps Stuff Safe) fournit des outils qui utilisent des ordinateurs locaux contrôlés par la bibliothèque pour sauvegarder l'accès à long terme du lecteur aux revues spécialisées basées sur la Toile. <http://lockss.stanford.edu/>

Long Term Preservation Study. Koninklijke Bibliotheek. (Régulièrement mise à jour.)

L'évolution des documents dans la Long Term Preservation Study entreprise dans le cadre du Project Depot van Nederlandse Elektronische Publicaties (DNEP) en association avec IBM. Un plan de recherche et une présentation des stratégies étudiées, comme le Preservation Layer Model, sont disponibles à partir du site <http://www.kb.nl/kb/ict/dea/ltp/ltp-en.html>

National Digital Information Infrastructure and Preservation Program (NDIIPP). Library of Congress. (Régulièrement mis à jour.)

Renferme des informations sur l'effort national d'organisation pour la préservation à long terme du contenu numérique en collaboration avec des représentants d'autres organismes fédéraux, de recherche, de bibliothèques et commerciaux. Liens avec de nombreuses publications du programme. <http://www.digitalpreservation.gov/ndiipp/>

Bibliothèque nationale d'Australie

Le site Web inclut des liens avec toute une série de programmes et de communications relatifs à la préservation numérique. <http://www.nla.gov.au/>

Collection Electronique de la Bibliothèque nationale du Canada. (Régulièrement mise à jour.)

Fournit l'accès aux versions archivées du matériel en ligne canadien. Inclut un lien avec les informations relatives à la collection électronique et à son histoire, ainsi qu'à des astuces pour l'archivage d'une publication en ligne. <http://collection.nic-bnc.ca/e-coll-e/index-e.htm>

NEDLIB: Networked European Deposit Library, Bibliothèque nationale des Pays-Bas. (Dernière mise à jour : février 2002.)

Page d'accueil de NEDLIB, un consortium de projet en collaboration dirigé par la Bibliothèque nationale des Pays-Bas (Koninklijke Bibliotheek) et qui inclut huit autres bibliothèques nationales

européennes, un service d'archives nationales et trois grands éditeurs. L'objectif principal était de trouver des manières de préserver l'accès aux publications numériques à la fois en ligne et hors connexion (format physique) en tant qu'infrastructure de base sur laquelle puisse être créée une bibliothèque de dépôt européenne mise en réseau. Les communications de travail techniques du projet sont disponibles en suivant le lien "Working Papers" sur la page d'accueil. <http://www.kb.nl/coop/nedlib/>

netarchive.dk. Denmark's Electronic Research Library. (Régulièrement mise à jour.)

Une initiative conjointe, qui associe la Royal Library, la State & University Library et le Centre for Internet Research, de l'Université d'Aarhus au Danemark. La longue étude de 2001-2002 étudiait la saisie et l'archivage de l'activité Internet danoise ayant trait aux élections municipales de 2001. <http://www.netarchive.dk>

Online Computer Library Center Inc. (OCLC) Digital and Preservation Resources. (Régulièrement mis à jour.)

L'OCLC est un service informatique et un organisme de recherche à but non lucratif dont le réseau et les services relient plus de 30.000 bibliothèques dans 65 pays et territoires. <http://www.oclc.org/digitalpreservation/>

Preservation of Scientific Serials: Three Current Examples. William Y. Arms, in: *The Journal of Electronic Publishing*, Volume 5, 2<sup>ème</sup> parution. (Date de création : décembre 1999.)

Etudie trois revues sous forme numérique : l'*ACM Digital Library*, l'*Internet RFC series*, et *D-Lib Magazine* et discutent des mesures qu'on peut prendre aujourd'hui pour préserver l'accès aux informations contenues dans ces revues. Les solutions proposées sont "en partie techniques, en partie organisationnelles". Propose trois niveaux de préservation : la préservation de la présentation ; la préservation de l'accès, en conservant à la fois le matériau sous-jacent et un système d'accès efficace ; et la préservation du contenu. <http://www.press.umich.edu/jep/05-02/arms.html>

Research Libraries Group (RLG). (Régulièrement mis à jour.)

La préservation est une des activités essentielles du RLG avec la conservation à long terme des matériaux de recherche numérique qui comprennent un domaine clé. <http://www.rlg.org/longterm/>

Portail des bibliothèques de l'UNESCO.

Le site Web inclut des liens avec de nombreux sites de bibliothèques importantes et des informations sur les projets de l'UNESCO. [http://portal.unesco.org/ci/ev.php?URL\\_ID=6513&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201&reload=1041937729](http://portal.unesco.org/ci/ev.php?URL_ID=6513&URL_DO=DO_TOPIC&URL_SECTION=201&reload=1041937729)

## **20.7 La préservation des matériaux d'archives**

Archival Preservation of Smithsonian Web Resources: Strategies, Principles, and Best Practices. Dollar Consulting. (Dernière mise à jour : 18 octobre 2001.)

Commandé par les Smithsonian Institution Archives, ce rapport fournit des directives sur la saisie, la gestion et la préservation à long terme des sites Web de la Smithsonian Institution. En incorporant un modèle de processus de durée de vie d'archives intégrées, il recommande les meilleures solutions, et propose des annexes sur la documentation des sites Web et un modèle de métadonnées de préservation. <http://www.si.edu/archives/archives/dollar%20report.html>

Conversion and Migration Criteria in Records Keeping Systems. Association for Information Management Professionals. (Régulièrement mis à jour.)

La norme proposée aborde les problèmes fondamentaux politiques, méthodologiques et techniques liés à la conversion et au passage d'un système de conservation d'archives à un autre indépendamment du format de la pièce d'archive, de sorte que ces systèmes garantissent le contexte, le contenu et la structure d'archives authentiques. <http://www.ama.org//publications/standards/workinprogress.cfm>

DAVID : Digitale Archivering in Vlaamse Instellingen en Diensten (Digital Archiving in Flemish Institutions and Administrations). Archives de la Ville d'Anvers. (Dernière mise à jour : 10 janvier 2003.)

Le projet DAVID est une collaboration entre les Archives de la Ville d'Anvers et l'ICRI qui a pour objectif de rechercher une longévité numérique dans un environnement gouvernemental. Il cherche à élaborer les meilleures solutions pour l'archivage des données concernant les élections et la population, des courriels et des sites Web. <http://www.antwerpen.be/david/>

Diffuse: Guide to Archiving. (Dernière mise à jour : mai 2002.)

Un guide d'archivage de données élaboré dans le cadre du projet Diffuse du programme de l'EU IST. Discute des nécessités essentielles, des liens avec les normes clés, des spécifications, des exemples des meilleures solutions et des rapports officiels sur l'archivage des enregistrements publics. [http://www.diffuse.org/archive\\_guide.html](http://www.diffuse.org/archive_guide.html)

Effective Records Management Project. Université de Glasgow.

Projet qui vise à produire des outils et des protocoles, ainsi qu'un système pilote pour la création et la diffusion des rapports du comité au sein de l'université. Un rapport final, publié au début de l'année 2002, est disponible en PDF sur les pages Web du projet. <http://www.gla.ac.uk/InfoStrat/ERM/>

Enduring Paradigm, Nex Opportunities: The Value of the Archival Perspective in the Digital Environment. Anne J. Gilliland-Swetland, dans : *CLIR Reports*. (Date de création : février 2000.)

Etudie l'utilité de l'optique de l'archivage en abordant des problèmes liés à la préservation de l'information numérique. <http://www.clir.org/pubs/reports/pub89/contents.html>

Guidelines for Electronic Records Management on State and Federal Agency Websites. Charles R. McClure; J. Timothy Sprehe.

Directives élaborées dans le cadre d'un projet de recherche sur la gestion des archives et les stratégies de préservation pour les informations électroniques contenues dans les sites Web de l'Etat (Etats-Unis) et de l'agence fédérale. <http://istweb.syr.edu/~mcclure/guidelines.html>

International Council on Archives. (Régulièrement mis à jour.) <http://www.ica.org/>

International Records Management Trust (IRMT). (Régulièrement mis à jour.)

Basé à Londres, l'organisme du IRMT a été fondé en 1989 pour aider les pays en développement à gérer leurs archives officielles. Les liens avec les projets parrainés par le IRMT et les rapports sont consultables sur ce site Web, y compris les ressources examinant les problèmes de gestion des archives électroniques. <http://www.irmt.org/index.html>

National Archives and Records Administration (NARA) (Etats-Unis). (Régulièrement mis à jour.) <http://www.archiv.gov/>

National Archives of Australia. (Régulièrement mis à jour.) <http://www.naa.gov.au/recordkeeping/preservation/summary.html>

National Archives of Singapore. (Régulièrement mis à jour.) <http://www.museum.org.sg/NAS/nas.shtml>

National Historical Publications and Records Commission (NHPRC). (Régulièrement mis à jour.)

Le NHPRC a établi une grande partie de sa recherche sur la préservation à long terme de l'accès aux archives électroniques. Le site Web fournit des liens avec les rapports du projet. La partie consacrée aux projets d'archives électroniques est particulièrement intéressante. <http://www.archives.gov/grants/index.html>

Public Record Office (PRO).

Le Public Record Office du Royaume-Uni vise à aider et promouvoir l'étude du passé en sélectionnant, préservant et fournissant l'accès aux archives publiques. Deux de ses projets de préservation numérique sont : l'EROS (Electronic Records from Office Systems) et le NDAD (National Digital Archive of Datasets-Royaume-Uni). <http://www.pro.gov.uk/>

Portail des Archives de l'UNESCO.

Le site Web inclut les liens avec de nombreux sites d'archives importants et avec des informations sur les projets de l'UNESCO. [http://portal.unesco.org/ci/ev.php?URL\\_ID=5761&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201&reload=1036751929](http://portal.unesco.org/ci/ev.php?URL_ID=5761&URL_DO=DO_TOPIC&URL_SECTION=201&reload=1036751929)

Victorian Electronic Records Strategy Project. Public Records Office, Victoria. (Dernière mise à jour : 31 mars 1999.)

Le projet vise à prouver la faisabilité de la saisie et de la préservation des archives électroniques ; et à fournir un ensemble de descriptions fonctionnelles pour l'archivage électronique. Les conclusions des projets, les descriptions fonctionnelles et une description générale du système de démonstration ont été publiés dans le Victorian Electronic Records Strategy Final Report. <http://www.prov.vic.gov.au/vers/published/final.htm>

## **20.8 La préservation des matériaux audio et audiovisuels**

Building a National Strategy for Digital Preservation: Issues in Digital Media Archiving, Council on Library and Information Resources (CLIR). (Date de création : avril 2002.)

Un ensemble de rapports et d'études commandés par la Library of Congress et le CLIR comme documentation du National Digital Information Infrastructure and Preservation Program. Les sujets traités englobent six domaines principaux qui présentent les problèmes de gestion de collection : les sites Web importants, les livres électroniques, les journaux électroniques, le son enregistré numériquement, le film numérique, et la télévision numérique. <http://www.clir.org/pubs/reports/pub106/contents.html>

The Care and Handling of Recorded Sound Materials. Bibliothèque nationale du Canada ; Gilles Saint-Laurent. (Dernière mise à jour : 15 février 2002.)

Fournit une explication de base de ce qu'est le son, comment il est enregistré, et identifie de nombreux supports différents d'enregistrement. Discute de la manipulation et de la préservation. <http://palimpsest.stanford.edu/byauth/st-laurent/care.html>

Digital Preservation of Moving Image Material? Howard Besser. (Date de création : 2001.)

Cet article décrit les changements de la technologie numérique qui surviennent dans les processus de production et de diffusion des matériaux de l'image animée. Indique deux changements de modèles probables pour la préservation de l'image animée : la gestion des oeuvres complètes c. contre celle du composant et l'artefact physique c. le contenu. Il aborde également les méthodes générales en matière de préservation numérique et les problèmes que rencontrent les archivistes de l'image animée. <http://www.gseis.ucla.edu/~howard/Papers/amia-longevity.html>

European Convention for the Protection of the Audiovisual Heritage and Protocol on the Protection of Television Productions. Conseil de l'Europe. (Date de création : 6 février 2001.)

Cette Convention, publiée par le Parlement européen, veille à la sauvegarde et à la préservation du patrimoine européen de l'image animée. Les signataires de l'accord ont l'obligation de mettre en place des mécanismes juridiques ou spontanés pour le dépôt des supports audiovisuels dans des gisements d'archivages désignés sur leur territoire. Le texte est rédigé en termes généraux pour que la législation puisse s'appliquer aux nouvelles formes électroniques et autres de l'expression audiovisuelle au fur et à mesure de leur création. [http://www.coe.int/t/e/cultural\\_cooperation/culture/Resources/Reference\\_texts/Conventions/econpataud.asp](http://www.coe.int/t/e/cultural_cooperation/culture/Resources/Reference_texts/Conventions/econpataud.asp)

PRESTO - IST-1999-2003: Key Links System Specification Document. Presto Consortium. (Dernière mise à jour : 26 juin 2001.)

Un rapport de 135 pages élaboré à partir des résultats d'une enquête sur la préservation des matériaux audiovisuels dans les Archives européennes de diffusion. Comporte des informations sur les besoins du système, les modernisations technologiques, les méthodologies en cours et les spécification de données pour l'archivage numérique du film, de l'audio et de la vidéo. <http://presto.joanneum.ac.at/Public/D32.pdf>

The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy. IASA Technical Committee. (Date de création : septembre 2001.)

Identifie les domaines du problème et recommande des solutions pour les archives audio et audiovisuelles, en comparant la situation idéale et la réalité. <http://www.iasa-web.org/iasa0013.htm>

UPF (Universal Preservation Format) Home Page

Parrainée par la WGBH Educational Foundation et créée en grande partie grâce à une subvention de la National Historical Publications and Records Commission, cette initiative préconise un format non tributaire d'une plate-forme qui aidera à rendre accessible un large éventail de types de données. L'UPF est défini comme "autodescriptif" car il inclut, dans ses métadonnées, toutes les spécifications techniques nécessaires pour créer et recréer des navigateurs de supports afin de pouvoir toujours accéder aux matériaux contenus. <http://info.wgbh.org/upf/>

## **20.9 La préservation des collections de données**

Archiving Scientific Data. Peter Buneman, Sanjev Khanna, Keishi Tajima, Wang-Chiew Tan. (Date de création : juin 2002.)

Décrit le développement d'un outil d'archivage pour les données XML qui permet de conserver tous les états précédents des données au fur et à mesure qu'elles se modifient avec le temps. Les descriptions de modification significatives, la récupération des versions spécifiques et l'historique sont pris en compte et soutenus. <http://db.cis.upenn.edu/Research/ki.html>

Arts and Humanities Data Service (AHDS). (Régulièrement mis à jour.)

L'AHDS est un service national (Royaume-Uni) qui collecte, décrit et préserve les ressources électroniques provenant de la recherche de l'enseignement des Lettres. L'un de ses objectifs est de développer des stratégies pour préserver le patrimoine culturel numérique. <http://ahds.ac.uk/>

Geophysical Data in Archeology: a guide to good practice. Armin Schmidt. (Date de création : 2002.)

Propose une introduction à la géophysique archéologique et à la diversité des données produites y compris les données de mesures brutes, les données exploitées et les dessins interprétatifs. Elle fournit aussi une introduction précieuse au stockage et à l'archivage des bases de données géophysiques. <http://ads.ahds.ac.uk/project/goodguides/geophys/>

Inter-University Consortium for Political and Social Research. (Régulièrement mis à jour.)

Au sein de l'Université du Michigan, l'ICPSR acquiert et préserve des données de sciences sociales pour le compte de 400 collèges et universités aux Etats-Unis et à l'étranger. Utilise des méthodes de transfert pour garantir un accès permanent aux données archivées. <http://www.icpsr.umich.edu/index.html>

Long Term Archiving of Digital Documents in Physics - Conference Report. Dr Arthur P. Smith.

Le rapport de la Conférence de l'IUPAP (International Union of Pure and Applied Physics) offre un résumé des débats sur le thème : "En quoi consistent des archives numériques ?" et répertorie les recommandations de la Conférence. [http://publish.aps.org/IUPAP/ltaddp\\_report.html](http://publish.aps.org/IUPAP/ltaddp_report.html)

NDAD: National Digital Archive of Datasets (Royaume-Uni). Le NDAD du London Computing Centre de l'Université de Londres abrite des données numériques archivées qui proviennent des ministères et des organismes du gouvernement du Royaume-Uni. <http://ndad.ulcc.ac.uk/>

Preservation of the Electronic Assets of a University. Oxford University Computing Services. T. Alex Reid. (Date de création : octobre 1997.)

Décrit la manière dont l'Université d'Oxford a abordé la gestion, le stockage et la préservation de ses biens électroniques. <http://users.ox.ac.uk/~alex/hfs-AXIS-paper.html>

## **20.10 La préservation de l'art numérique**

Archiving the Avant Garde: Documenting and Preserving Variable Media Art. Berkeley Art Museum and Pacific Film Archive. (Date de création : 2001.)

Projet mené en collaboration pour développer, documenter et propager des stratégies en vue de décrire et de préserver des formes d'art faisant appel à divers médias et intermédias non traditionnels, comme l'art de la performance, de l'installation, l'art conceptuel et l'art numérique. [http://www.bampfa.berkeley.edu/ciao/avant\\_garde.html](http://www.bampfa.berkeley.edu/ciao/avant_garde.html)

Longevity of Electronic Art. Howard Besser. (Date de création : février 2001.)

Met en lumière les problèmes de la préservation de l'art électronique, en signalant les caractéristiques particulières des oeuvres d'art électroniques qui constituent un défi pour la préservation et propose des stratégies pratiques pour préserver l'art électronique. <http://www.gseis.ucla.edu/~howard/Papers/elect-art-longevity.html>

Rhizome.org: The New Media Art Resource

Rhizome.org est une organisation à but non lucratif qui vise à préserver l'art électronique. <http://rhizome.org/info/index.php>

Variable Media Initiative. Guggenheim Museum.

Cherche à identifier des stratégies approuvées par les artistes pour préserver les créations qui utilisent des médias variables (l'installation, la performance, l'interactif, le numérique). Les artistes sont encouragés à définir leur oeuvre indépendamment du médium et à fournir des directives sur la manière dont elle peut être remaniée dans des nouveaux formats. <http://www.guggenheim.org/variablemedia/>

## **20.11 La préservation du courriel**

Archiving E-Mails. Filip Boudrez; Sofie Van den Eynde. (Date de création : septembre 1999.)

Un rapport du projet flamand DAVID qui examine en détail les problèmes juridiques et techniques liés à la préservation des archives de courriel. <http://www.antwerpen.be/david/teksten/Report4.pdf>

E-Mail and Potential Loss to Future Archives and Scholarship or The Dog that Didn't Bark. Susan Lukesh, in: *First Monday*. (Date de création : août 2002.)

Cet article discute de l'importance de la communication informelle et de la manière dont elle est de plus en plus créée dans des formats électroniques qu'il faut activement préserver. Susan Lukesh préconise une action de la part des archivistes, des fabricants de logiciels, des institutions publiques et des créateurs, en particulier des savants, pour aider à la préservation du courriel. [http://www.firstmonday.dk/issues/issue4\\_9/lukesh/](http://www.firstmonday.dk/issues/issue4_9/lukesh/)

E-Mail-XML Demonstrator : Technical Description. Testbed Digitale Bewaring. (Date de création : octobre 2002.)

Ce rapport décrit le logiciel prototype développé par le projet Dutch Testbed Digitale Bewaring lors de ses recherches sur la préservation à long terme des messages du courriel. La solution est fondée sur la personnalisation du Microsoft Outlook pour permettre une communication avec un serveur central responsable de la collection de données, de la conversion et de l'archivage à la fois des messages et des métadonnées en XML. <http://www.digitaleduurzaamheid.nl/bibliotheek/docs/email-demo-en.pdf>

Strategies for Capturing and Managing Emails as records and as Organisational Assets. Adrian Cunningham. (Date de création : 18 avril 2002.) [http://www.naa.gov.au/recordkeeping/noticeboard/emails\\_as\\_records\\_files/frame.htm](http://www.naa.gov.au/recordkeeping/noticeboard/emails_as_records_files/frame.htm)

## **20.12 La préservation des collections de prêtirages électroniques**

E-print Services and Long-term Access to the Record of Scholarly and Scientific Research. Michael Day, in *Ariadne*. (Date de création : 22 juin 2001.)

Examine certains des problèmes que pose la préservation à long terme aux services de prêtirage électronique. Aborde également certains aspects clés comme la responsabilité de préservation et l'authenticité. <http://www.ariadne.ac.uk/issue28/metadata/>

Setting Up An Institutional E-print Archive. Michael Gardner, John MacColl, Stephen Pinfield, in: *Ariadne*. (Date de création : 16 avril 2002.)

Fondé sur plusieurs expériences de mise en place de serveurs pilotes de prêtirage électronique dans les Universités d'Edimbourg et de Nottingham ; fait état de plusieurs problèmes pratiques, dont les types et les formats de document, les procédures de présentation, les normes des métadonnées et les problèmes de préservation numérique. <http://www.ariadne.ac.uk/issue31/eprint-archives>

SHERPA: Securing a hybrid environment for research, preservation and access. (Dernière mise à jour : 2002.)

Le SHERPA est un projet qui s'échelonne sur trois ans fondé par le JICS en vue de créer des "archives de prêtirages électroniques" pour les principales institutions de recherche du Royaume-Uni. Les archives se conformeront au protocole de collecte de métadonnées de l'Open Archives Initiative et envisagent pour la préservation numérique d'explorer la voie de l'application du modèle de référence de l'OAIS. <http://www.sherpa.ac.uk>

## **20.13 La préservation des matériaux numériques de format physique**

Bits is Bits: Pitfalls in Digital Reformatting. Walt Crawford. (Date de création : mai 1999.)

Cet article décrit certaines des entraves au reformatage des matériaux numériques, comme la technologie de protection de copie, les dépendances logicielles et matérielles, et le cryptage.

*American Libraries*, vol. 30, n° 5 (05/99)

CD-R and CD-RW Questions and Answers. Optical Storage Technology Association (OSTA) (Date de création : 2001.)

Cette série de pages, dues l'Optical Storage Technology Association, embrasse un certain nombre de sujets concernant les supports CD-ROM et les CR-RW, y compris quelques définitions de terme, la longévité des supports, la manipulation, l'étiquetage, la vitesse et la qualité. <http://www.osta.org/technology/cdqa.htm>

Farewell my Floppy: a strategy for migration of digital information. Deborah Woodyard. (Dernière mise à jour : 29 avril 1998.)

Cet article décrit une enquête sur les matériaux de la collection de la Bibliothèque nationale d'Australie stockés sur disquette et rend compte des aspects pratiques du transfert de disquettes souples sur CD-ROM. <http://www.nla.gov.au/nla/staffpaper/valadw.html>

Mapping Functionality of Off-line Archiving and Provision Systems to OAIS. Jorg Berkemeyer. Die Deutsche Bibliothek. (Date de création : janvier 1999.)

Commente la préservation des matériaux numériques de format physique entreprise par les bibliothèques nationales et dans le contexte du modèle de référence de l'OAIS. <http://www.kb.nl/coop/nedlib/meetings/frankfurt/GEN-232.doc>

## **20.14 La numérisation**

Colorado Digitization Project Digital Toolbox. (Mis à jour régulièrement.)

La Digital Toolbox est conçue pour aider les administrateurs à se poser les bonnes questions dans les étapes initiales de planification d'un projet numérique. Fournit des informations sur les aspects techniques de la numérisation. <http://www.cdpheritage.org/resource/index.html>

Guides to Quality in Visual Resource Imaging. Donald P. D'Amato, Franziska Frey, Linda Serenson Colet, Don Williams. (Date de création : juillet 2000.)

Cinq guides écrit conjointement avec la Digital Library Federation, le CLIR et le RLG pour identifier les technologies de l'image et des solutions pour les ressources visuelles. Des informations pratiques sur la planification d'un projet, le choix d'un numériseur, les facteurs nuisant à la qualité de l'image, la mesure de la qualité de l'image, et les formats de fichier pour les fichiers originaux. <http://www.rlg.org/visguides/>

Handbook for Digital Projects: A Management Tool for Preservation and Access. Maxine K. Sitts (Ed.). (Date de création : décembre 2000.)

Une ressource Web fournissant des informations sur les problèmes relatifs à la conversion numérique des matériaux de collection. Des articles écrits par de nombreux présentateurs de la School for Scanning fournissent des informations sur la gestion et le choix du projet, sur la longévité numérique, le copyright et des considérations techniques. <http://www.nedcc.org/digital/dighome.htm>

nof-digitise Technical Standards and Guidelines. People's Network Development Team. (Régulièrement mis à jour.)

Un guide technique des projets de numérisation mis au point par UKOLN and Resource: The Council for Museums, Archives & Libraries for the New Opportunities Fund. Adopte une approche de durée de vie et esquisse les étapes successives de la création, du développement, de la gestion, de l'accès et de la réutilisation de l'information numérique. <http://www.peoplesnetwork.gov.uk/content/technical.asp>

Selection Criteria for Digital Imaging. Columbia University Library. (Dernière mise à jour : 14 janvier 2001.) <http://www.cc.columbia.edu/cu/libraries/digital/criteria.html>

The Society for Imaging Science and Technology. (Régulièrement mis à jour.)

Une association à but non lucratif dont l'objectif est d'informer ses membres des derniers développements scientifiques et technologiques dans le domaine de l'image. <http://www.imaging.org/>

Technical Advisory Service for Images (TASI). (Régulièrement mis à jour.) Le TASI est un service créé pour conseiller et aider la communauté universitaire du Royaume-Uni dans la création

numérique, le stockage et la livraison des informations liées à l'image. Fournit des informations sur la préservation de l'accès aux images numériques. <http://www.tasi.ac.uk/>

## **20.15 Le dépôt légal et le dépôt spontané**

Dépôt légal et numérotations. Bibliothèque nationale de France. (Régulièrement mis à jour.)

Version actualisée des pages Web du dépôt légal de la Bibliothèque nationale de France. En même temps qu'elle fournit des informations générales sur la procédure obligatoire, elle fournit des liens avec la législation actuelle et présente les recommandations du Conseil scientifique du dépôt légal qui préconisent que le dépôt soit étendu aux publications en ligne. <http://www.bnf.fr/pages/zNavigat/frame/infopro.htm>

Legal Deposit from the Internet in Denmark: Experiences with the Law from 1997 and the Need for Adjustments. (Date de création : juin 2001.)

In la conférence *Papers from the Preserving the Present for the Future: Strategies* qui s'est déroulée dans la Royal Library à Copenhague. <http://www.deflink.dk/eng/arkiv/dokumenter2.asp?id=695>

Legal Deposit. Bibliothèque nationale d'Ecosse. (Régulièrement mis à jour.)

Liens avec une documentation concernant le dépôt des publications du Royaume-Uni non imprimées et une Revised Version of the Code of Practice for the Voluntary Deposit of Non-Print Publications avec les notes explicatives s'y rapportant. Concerne le dépôt des publications électroniques à la fois en ligne et hors ligne, les dernières faisant l'objet d'un essai de dépôt expérimental. <http://www.nls.uk/professional/legaldeposit/index/html>

Management of Networked Electronic Publications: A Table of Status in Various Countries. Elizabeth Martin. (Dernière mise à jour : novembre 2001.)

Une comparaison entre 16 bibliothèques nationales à propos de la législation et des dispositifs du dépôt, de l'approche et de la politique, des plans, des négociations avec les éditeurs, des dispositifs et de l'installation de l'accès pour les publications électroniques mises en réseau. <http://www.nlc-bnc.ca/obj/r7/f2r7-100-f.pdf>

Disponible en français à <http://nlc-bnc.ca/obj/R7/f2/r7-100-f.pdf>

A standard for the legal deposit of online publications. Giovanni Bergamin. (Date de création : 4 juin 1999.)

Résumé en anglais ; résumé en italien disponible à :

<http://www.aib.it/aib/commiss/cnur/dliberga.htm>

<http://www.aib.it/aib/commiss/cnur/dleberga.htm>

Statement on the Development and Establishment of Codes of Practice for the Voluntary Deposit of Electronic Publications. Conference of European National Librarians (Date de création : 2000.)

Déclaration officielle commune de la conférence des Bibliothèques nationales européennes et de la Fédération des éditeurs européens. Compris un projet de Code de Pratique, destiné à faciliter la rédaction de dispositions de dépôt spontané approuvées localement. <http://minos.bl.uk/gabriel/fep/>

## **20.16 Les métadonnées**

Digital Libraries: Metadata Resources. International Federation of Library Associations and Institutions. (Dernière mise à jour : 22 septembre 1999.)

Relie à de nombreux articles et sites concernant la documentation et les normes des données.  
<http://www.ifla.org/II/metadata.htm>

Meta Matters. Bibliothèque nationale d'Australie.

Ce site Web est destiné à aider les fournisseurs de contenu sur la Toile à améliorer l'efficacité de la recherche de ressources d'information en utilisant les normes des métadonnées.  
<http://www.nla.gov.au/meta/>

Metadata Encoding & Transmission Standards (METS), Library of Congress (Date de création : 14 juin 2001.)

Site Web officiel du schéma METS XML pour le codage des métadonnées descriptives, administratives et structurelles. <http://www.loc.gov/standards/mets/>

Preservation Metadata and Digital Continuity. Steve Knight, in: *DigiCULT. Info Newsletter*. (Date de création : février 2003.)

Décrit de façon générale le programme de préservation numérique de la Bibliothèque nationale de Nouvelle-Zélande et le développement d'un schéma de métadonnées de préservation.  
[http://data.digicult.info/download/digicult\\_info3\\_low.pdf](http://data.digicult.info/download/digicult_info3_low.pdf)

UKOLN Metadata. Michael Day. (Régulièrement mis à jour.)

Un site de métadonnées générales qui fournit des liens avec des projets, des initiatives, des registres et des ressources, y compris des outils logiciels pour manipuler les métadonnées, et un glossaire.  
<http://www.ukoln.ac.uk/metadata/>

## **20.17 Les normes**

Digital Library Standards. University of California Libraries. (Régulièrement mis à jour.)

Fournit des liens avec des ressources concernant une gamme de normes pour bibliothèque numérique. <http://www.sunsite.berkeley.edu/Info/standards.html>

National Information Standards Organisation - NISO. (Régulièrement mis à jour.)

Développe et fait la promotion des normes techniques internationales utilisées dans les services d'information. <http://www.niso.org/>

PDF-Archive Project (PDF/A). Association for Information and Image Management, International (AIIM, International).

Une action commune de l'Association for Suppliers of Printing, Publishing and Converting Technologies (NPES) et de l'AIM, International, pour développer une norme internationale qui définisse l'utilisation de l'Adobe's Portable Document Format (PDF) pour l'archivage et la préservation des documents électroniques. Le projet fera appel à des documents multipage qui

présentent des combinaisons de texte et de graphiques et réclamera des appareils de lecture pour restituer des documents archivés. <http://www.aiim.org/standards.asp?ID=25013>

Standards for Libraries. Bibliothèque nationale d'Australie.

Ce site fournit des liens avec des informations sur la bibliothèque et les normes connexes, avec des listes de normes et les principaux organismes de normes. <http://www.nla.gov.au/services/standard3.html>

W3C World Wide Web Consortium. (Régulièrement mis à jour.)

Le W3C, un organisme professionnel qui vise à exploiter au maximum les potentialités de la Toile en développant des protocoles communs. <http://www.w3.org/>

XML for Digital Preservation: XML Implementation Options for E-Mails. Maureen Potter. (Date de création : 11 octobre 2002.)

Un rapport sur les progrès accomplis par la Digital Preservation Testbed (Testbed Digital Bewaring) (Pays-Bas) dans l'utilisation du langage XML comme méthode de préservation. <http://www.digitaleduurzaamheid.nl/bibliotheek/docs/email-xml-imp.pdf>

## **20.18 Quelques outils intéressants**

The Computer History Simulation Project. (Régulièrement mis à jour.)

Un collectif informel basé sur l'Internet de personnes qui s'intéressent à la restauration par simulation de systèmes de matériel et de logiciels informatiques historiquement significatifs. <http://simh.trailing-edge.com/>

My File Formats. (Régulièrement mis à jour.)

Un site Web qui comprend des informations sur plus de 1.000 formats de fichier. <http://myfileformats.com/?old=manufacturers&truespace=.com.html>

Software Archaeology, Andrew Hunt, David Thomas, in *IEEE Software*, vol. 19, n° 2 (mars/avril 2002).

Un article court qui décrit les problèmes que pose la compréhension du code du logiciel avec peu ou pas de documentation. Il s'achève par quelques propositions sur la manière dont les développeurs actuels pourraient faciliter le fonctionnement du code à l'avenir.

ISSN : 0740-7459

Windows Desktop Product Life Cycle Support and Availability Policies for Businesses. (Régulièrement mis à jour.)

Un article qui trace les grandes lignes de la politique de Microsoft en faveur d'un soutien permanent de ses produits commerciaux de bureau avec des échelles de temps et des détails sur certains produits spécifiques. <http://www.microsoft.com/windows/lifecycle/default.msp>

## **21. Index**

Prendre des responsabilités

Accès

Barrières d'accès

Conditions d'accès

Accessibilité

Transparence

Préconisation

Appréciation

Stockage d'archivage

Matériaux audiovisuels

Authenticité

Sauvegarde

Compatibilité en amont

Modèles d'entreprise

Modèles centralisés

Certification

Caractéristiques des programmes fiables

Collaboration

Accords de collecte

Développement de collection

Communauté

Programmes exhaustifs

Collecte exhaustive c. collecte sélective

Objets conceptuels

Plan d'éventualités

Continuité

Contrôle

Coopération

Coûts

Créateurs, voir Producteurs

Diversité culturelle

Supports de données

Dépendance des données

Extraction de données

Intégrité des données

Protection des données

Récupération des données

Stockage des données

Transfert de données

Bases de données

Ensembles de données

Prise de décisions

Affirmer ses responsabilités

Dépôt

Objets dérivés

"Communauté désignée"

Patrimoine numérique

Numérisation

Plan de anticatastrophe

Programmes répartis

Division du travail

Documentation

"Codage durable"

Courriels

Emulation

Encapsulation

Equipement

Elements essentiels

Compétence

Mécanismes à sécurité intégrée

Formats de fichier

Identification de fichier

Durabilité financière

Fonctions des programmes de préservation

Bailleurs de fonds

Collecte

Patrimoine

Institutions du patrimoine

Récolte

Fichier d'image

Partage de l'information

Ensemble cohérent d'information

Ingestion

Internet

Collections héritées

Dépôt légal

Cadre légal

Niveaux de responsabilité

Objets logiques

Pertes, niveaux acceptables de

Supports magnétiques

Gestion des programmes de préservation

Moyens d'accès

Métadonnées

Passage

Programmes minimaux

Stratégies multiples

Programmes non exhaustifs

Méthodes non numériques

Normalisation

OAIS (Open Archival Information Systems Reference Model)

Supports optiques

Structures organisationnelles

Viabilité organisationnelle

Partenaires

Approche de "performance"

Identifiants permanents

Objets physiques

Planification

Politique

Métadonnées de préservation

Programmes de préservation

Principes (résumé)

Impression

Etablir les priorités

Respect de la vie privée et confidentialité

Producteurs

Editeurs, voir Producteurs

Techniques de transfert à flux poussé

Techniques de transfert à flux tiré

Redondance

Programmes fiables

Métadonnées de découverte de ressources

Ressources

Responsabilités

Formats de fichier restreint

Droits

Gestion du risque

Rôles et responsabilités

"Lieux sûrs"

Sélection

Collecte sélective c. collecte exhaustive

Prestataires de service

Partage des coûts

Propriétés significatives, voir Eléments essentiels

Logiciel

Fabricant de logiciels et de matériel

Dépendance du logiciel

Réagencement du logiciel

Spécifications

Recruter du personnel

Parties prenantes

Normes

Programmes autonomes

Points de départ

Spécifications de stockage

Stratégies

Plans de succession

Sécurité du système

Adéquation entre techniques et procédures

Infrastructure technique

Préservation de technologie

Terminologie

Menaces

Transfert

Programme crédible

Types de patrimoine numérique

Campagne du Patrimoine numérique de l'UNESCO

Projet de charte de l'UNESCO pour la préservation du Patrimoine numérique

Universités (en tant que gardiennes du patrimoine numérique)

Utilisateurs

UVC ("Universal Virtual Computer")

Visualiseurs

Travailler ensemble