

[*Traduction des passages sélectionnés de Part IV, Methods of Appraisal and Preservation :*
Introduction, p. 161-166
Research Findings, p. 173-183
Addressing the Research Questions, p. 183-192
Towards Guidelines for Preserving Authentic Digital Records, p. 192]

Partie IV

Méthodes d'évaluation et de conservation

Rapport de l'équipe thématique 3

par

Yvette Hackett, Library and Archives Canada

Introduction¹

Parmi le très large éventail de sujets de recherche abordés par la deuxième phase du projet InterPARES, l'évaluation et la conservation des documents sont considérées par beaucoup comme le cœur de la profession d'archiviste. L'évaluation est cette fonction clé par laquelle la décision de l'archiviste détermine quels documents primaires seront disponibles dans le futur pour soutenir des procédures légales, la recherche historique, la généalogie – questions d'identité, de culture, d'histoire et de droits.

Quant aux décisions de conservation, elles peuvent garantir la survie sur le long terme des documents sélectionnés, les conserver d'une manière qui leur enlève toute crédibilité ou toute utilité futures ou encore les détruire par erreur ou omission, annulant ainsi les effets de l'évaluation.

Contexte et mission

En dépit de l'importance de l'évaluation et de la conservation pour les archivistes, le thème 3 semble avoir eu plus de difficultés que les autres à définir le périmètre de ses activités et à se fixer un plan de travail. La proposition de recherche initiale prévoyait que l'unité de recherche thématique 3 commencerait par travailler en s'appuyant sur deux outils élaborés par InterPARES 1. Dans un premier temps, elle devait fusionner le *Model of Selection Function*², de l'équipe Évaluation, et le *Model of Preservation Function*³, de l'équipe Conservation, avec le modèle *Manage Archival Funds*⁴ élaboré dans le cadre du projet UBC-MAS. Le « grand modèle unifié », comme il avait été initialement baptisé, qui en aurait résulté, aurait représenté le cycle de vie complet du document d'archives, depuis sa production initiale par le producteur jusqu'à sa conservation et son accès sur le long terme, une fois versé au service d'archives compétent. Le modèle achevé, des expériences de prototypage auraient été

¹ L'auteur remercie tous les membres de l'équipe thématique 3 pour leur contribution à l'établissement de ce rapport. L'auteur tient à remercier plus particulièrement Luciana Duranti, Ken Thibodeau et Randy Preston pour leur contribution au texte et leurs conseils éditoriaux. Toute erreur de représentation ou omission serait de la seule responsabilité de l'auteur.

² Voir Appraisal Task Force (2001). « Appendix 4 : A Model of the Selection Function », in *The Long-term Preservation of Authentic Electronic Records : Findings of the InterPARES Project*, Luciana Duranti, éd. (San Miniato, Italie : Archilab, 2005, 239-252. Une version en ligne est disponible à l'adresse suivante : http://www.interpares.org/display_file.cfm?doc=ip1_aptf_model.pdf.

³ Voir Preservation Task Force (2002). « Appendix 5 : A Model of the Preservation Function », version 6.0, *ibid.*, 253-292. Une version en ligne est disponible à l'adresse suivante : http://www.interpares.org/display_file.cfm?doc=ip1_ptf_model.pdf.

⁴ Voir « Appendix B: Activity Models », in Luciana Duranti, Terry Eastwood et Heather MacNeil, *Preservation of the Integrity of Electronic Records*, (Dordrecht : Kluwer Academic Publishers, 2002). Une version en ligne est disponible à l'adresse suivante : <http://www.interpares.org/UBCProject/a-Of.htm>.

conduites afin de montrer comment les fonctions décrites dans le modèle pourraient être intégrées à des applications logicielles.

Du point de vue de l'organisation de l'équipe, il s'est avéré difficile de faire travailler ensemble les modélisateurs aguerris qui avaient participé à InterPARES 1 et poursuivaient en fait un travail initié en 1999, et les nouvelles recrues d'InterPARES 2, dont beaucoup ne connaissaient pas grand chose aux activités archivistiques ni aux subtilités de la méthodologie de modélisation adoptée par le projet. Ce problème a été résolu en juin 2003, avec la constitution d'une unité de recherche distincte, l'équipe transversale Modélisation. La décision prise en février 2004 de créer un modèle supplémentaire reposant sur le concept de *records continuum* et non sur celui de cycle de vie a éloigné un peu plus les chercheurs du thème 3 des travaux de modélisation⁵.

Les chercheurs du thème 3 ont également renoncé à l'idée d'utiliser, comme cela avait été envisagé, les études de cas d'InterPARES 1 pour valider les modèles d'InterPARES 1 préalablement à toute activité de prototypage de système. Les modifications à apporter aux modèles d'origine d'InterPARES 1 pour créer le modèle de Chaîne archivistique (*Chain of Preservation*, COP) indiquaient que les deux modèles originaux étaient trop dépassés pour être utiles aux travaux actuels. Quant au modèle d'Archivage orienté métier (*Business-driven Recordkeeping*, BDR), aucune analyse ne pouvait être engagée tant qu'il n'était pas achevé, ce qui, étant donné son lancement tardif, ne pourrait intervenir que plus tard dans le projet.

En outre, alors que les études de cas d'InterPARES 2 étaient en cours de sélection, il est apparu que les études de cas d'InterPARES 1 ne pouvaient être utilisées pour engager, en attendant, les travaux du thème 3. Les environnements d'archivage envisagés pour les travaux d'InterPARES 2 étaient en effet totalement différents de ceux étudiés dans InterPARES 1, même pour la sphère de l'administration. Comme pour le groupe de modélisation, seul un nombre réduit de membres du thème 3 avait participé à InterPARES 1 et connaissait bien les études de cas antérieures. Il est rapidement apparu que la « mise à niveau » de l'équipe exigerait un travail considérable, tout en ne présentant que peu d'intérêt scientifique relativement aux systèmes dynamiques, interactifs et expérimentiels étudiés par InterPARES 2.

Au moment où le rapport de mi-parcours était diffusé au printemps 2004⁶, il ne restait quasiment plus rien au thème 3 de sa feuille de route initiale :

⁵ Voir le rapport de l'équipe transversale Modélisation : http://www.interpares.org/display_file.cfm?doc=ip2_book_part_5_modeling_task_force.pdf. Pour les parties traduites en français, voir Partie V – Modéliser la production, la maintenance et la conservation des documents d'archives numériques.

⁶ Voir : http://www.interpares.org/ip2/ip2_midterm_progress.cfm.

... les activités planifiées au départ pour les deux premières années de travail des équipes thématiques 1 et 3 ont été réaffectées à une nouvelle unité de recherche, l'équipe transversale Modélisation⁷...

Quant aux outils et documents que devait produire l'équipe - « prototypes de systèmes d'évaluation et de conservation, modèles d'activité, et principes directeurs à l'usage des services et organismes en charge de la conservation des documents sur long terme⁸ » -, seul le concept de principes directeurs a été maintenu.

Enfin, l'équipe thématique 3 s'est heurtée à une dernière difficulté. Malgré le nombre élevé de participants, le projet ne comptait qu'une poignée d'archivistes. Il y avait neuf groupes de travail⁹, regroupés en trois unités de recherche sectorielles ou en trois unités de recherche thématiques et quatre unités de recherche transversales, qui requéraient la participation d'archivistes expérimentés – petits et grands services d'archives, entreprises et organismes publics, institutions indépendantes et institutions rattachées à de grandes organisations, et des archivistes ayant une bonne connaissance des documents du secteur public et du secteur privé, et des domaines artistiques, scientifiques et de l'administration. Enfin, les vingt-trois études de cas et les onze études générales requerraient également la participation d'archivistes. De fait, les archivistes n'étaient pas en nombre suffisant et ce problème a été exacerbé par la difficulté de certaines institutions archivistiques participantes à maintenir une représentation constante tout au long du projet.

Au printemps 2005, alors que plusieurs études de cas et études générales étaient achevées et qu'il ne manquait plus au modèle de Chaîne archivistique que son document de présentation, les membres du thème 3 ont fini par admettre que, dans InterPARES 2 comme dans la vie, l'évaluation et la conservation viennent en dernier. L'heure était venue pour le thème 3 de commencer son travail.

La proposition de recherche d'InterPARES 2 contenait une description de la fonction d'évaluation, insistant sur les différences entre l'évaluation de documents d'archives numériques et l'évaluation de documents papier :

L'évaluation juge de la valeur continue des documents mais elle rassemble également des preuves pour la présomption de leur authenticité et identifie les composants ou les objets numériques qui

⁷ Luciana Duranti, (2004). « InterPARES 2 Project Midterm Report to the Social Sciences and Humanities Research Council ». MCRI Grant no. 412-2001-1003, 4 (non publié).

⁸ Site du projet InterPARES 2, page du thème 3. Voir : http://www.interpares.org/ip2/ip2_domain3.cfm.

⁹ Pour un résumé de l'organisation intellectuelle du projet, voir : http://www.interpares.org/ip2/ip2_intellectual_organization.cfm.

doivent être stockés et reproduits en vue de garantir la conservation de documents authentiques. L'évaluation détermine également la faisabilité de la conservation d'un ensemble donné de documents électroniques en fonction des capacités de conservation actuelles et futures du service d'archives compétent¹⁰.

Cette définition reflète les conclusions d'InterPARES 1, qui avaient mis en lumière la nécessité de repenser la méthodologie traditionnelle d'évaluation afin de prendre en compte les spécificités de l'environnement numérique, et ce, à trois égards.

Premièrement, étant donné la facilité avec laquelle les documents numériques peuvent être endommagés ou modifiés, il est indispensable de documenter les preuves à l'appui de l'authenticité des documents. Avec les documents analogiques, l'authenticité est souvent présumée, en particulier lorsque les documents sont acquis directement auprès du producteur, puisque dans ce cas la provenance est connue et que la chaîne archivistique est ininterrompue du producteur à l'institution en charge de la conservation.

Deuxièmement, la définition ci-dessus insiste sur le concept de « composants numériques », les divers morceaux qui doivent être identifiés et préservés pour reproduire le document complet. Avec les documents analogiques, dans la grande majorité des cas, le contenu du document et sa structure ne peuvent être modifiés ou séparés du support – les différents composants du document sont inextricablement liés. De nombreuses informations sur le contexte de production du document peuvent également être intégrées à ce tout immuable, notamment en portant sur le document des numéros de dossiers ou des cotes, et par la présence ou l'absence de caractères de forme.

Il y a, enfin, le fait que des options de conservation complexes doivent désormais être évaluées et leur coût estimé avant qu'un service ou une institution de conservation ne puisse s'engager de manière réaliste à conserver des documents numériques sur le long terme. L'éventail des politiques de conservation numérique va bien au-delà de l'adoption de dossiers et de boîtes en carton neutre, même si l'évolution des documents analogiques n'est pas allée sans un cortège de formats de plus en plus difficiles à préserver – du papier acide aux négatifs et aux pellicules contenant des nitrates, en passant par le papier thermique pour fax et les bandes magnétiques audio.

En tant qu'unité de recherche en charge d'explorer les méthodes d'évaluation et de conservation des

¹⁰ Luciana Duranti (2001). « International Research on Permanent Authentic Records in Electronic Systems (InterPARES) : Experiential, Interactive and Dynamic Records », SSHRC MCRI InterPARES 2 Project Proposal, 412-2001, 1.1-12. Voir : http://www.interpares.org/display_file.cfm?doc=ip2_detailed_proposal.pdf.

documents d'archives numériques, il a été demandé au thème 3 d'évaluer si les concepts développés par l'équipe Évaluation et par l'équipe Conservation d'InterPARES 1 pouvaient être pertinents pour les environnements numériques étudiés par InterPARES 2. Ces environnements dynamiques, interactifs et/ou expérientiels sont en effet susceptibles de produire des documents sans équivalent évident dans l'univers papier.

Le rapport final de l'équipe Évaluation d'InterPARES 1 avançait un certain nombre de suggestions concernant la pratique de l'évaluation en environnement numérique¹¹. À partir de l'analyse d'études de cas portant pour l'essentiel sur des bases de données et des systèmes de gestion électronique des documents et des documents d'archives, et sur la base d'un modèle de la fonction d'évaluation, l'équipe Évaluation est parvenue aux conclusions suivantes :

- Il est préférable de conduire l'évaluation des documents numériques lorsque ceux-ci sont toujours utilisés par leur producteur. L'évaluation des documents numériques au cours des premières phases de leur cycle de vie permet à l'archiviste de disposer d'informations plus complètes sur le rôle opérationnel des documents dans l'organisation du producteur et fournit des informations techniques sur la façon dont l'application crée et maintient les archives courantes et intermédiaires. On est en droit d'espérer que l'identification précoce des documents présentant une valeur archivistique améliorera leurs chances de ne pas être détruits de manière accidentelle ou de ne pas sombrer dans une obsolescence technologique irrémédiable.
- La nature du support des documents a une incidence sur le processus d'évaluation mais non sur l'évaluation en tant que telle. L'identification formelle d'indicateurs d'authenticité¹², l'évaluation des politiques de conservation et le contrôle régulier des résultats de l'évaluation constituent de nouvelles tâches ou des tâches élargies du processus d'évaluation.
- Effectuer un suivi de la décision découlant de l'évaluation pour confirmer la valeur archivistique des documents sélectionnés est indispensable dans l'environnement numérique. L'idée que les décisions de conservation doivent être réexaminées à intervalles réguliers se justifie d'autant plus lorsque les évaluations sont conduites, comme suggéré plus haut, très en

¹¹ Voir Heather MacNeil et al. « Part One – Establishing and Maintaining Trust in Electronic Records : Authenticity Task Force Report », in Duranti, *Long-term Preservation*, op. cit., 19-65. Une version électronique est disponible à l'adresse suivante : http://www.interpares.org/book/interpares_book_d_part1.pdf.

¹² Voir Authenticity Task Force (2002). « Appendix 2 : Requirements for Assessing and Maintaining the Authenticity of Electronic Records », *ibid.*, 204-219. Une version électronique est disponible à l'adresse suivante : http://www.interpares.org/book/interpares_book_k_app02.pdf.

amont du versement des archives définitives à l'institution de conservation compétente, et aussi pour suivre le rythme du changement technologique.

- Les informations réunies au cours de l'évaluation doivent être « emballées » et transmises pour aider au suivi, au versement, au traitement, à la description, à la conservation et à l'exploitation future des documents. L'automatisation de tous les aspects du travail archivistique, y compris l'évaluation, facilitera grandement cette réutilisation continue, dans des phases ultérieures, d'informations réunies au cours du processus d'évaluation.

S'appuyant sur une méthodologie similaire, l'équipe Conservation d'InterPARES 1 présentait les conclusions suivantes dans son rapport final¹³ :

- Il n'est pas possible de conserver un document numérique : seule la capacité de le reproduire peut être conservée. À l'instar des conclusions relatives à l'évaluation, ce constat souligne l'importance de la notion de « composants » dès lors qu'il est question de documents numériques. Cette notion n'est pas absente des technologies analogiques – on pense par exemple au négatif et au tirage en photographie, ou au négatif et à l'image positive, la bande sonore optique et/ou magnétique, la copie composite et les chutes dans le domaine de l'image animée. Avec les documents numériques, ce système de composants acquiert une complexité inédite, qui exige d'être particulièrement attentif aux multiples dépendances pouvant être occasionnées par le matériel, les systèmes d'exploitation et les logiciels.
- Les composants intellectuels et physiques des documents numériques ne coïncident pas nécessairement ; un composant numérique est distinct d'un élément de forme. Par exemple, le contenu d'un document peut comporter à la fois du texte contenu dans un fichier de traitement de texte et un tableau produit par un tableur. Techniquement, le fichier peut ne contenir qu'un lien vers le fichier du tableau, dont l'affichage peut lui-même dépendre du tableur plutôt que du logiciel de traitement de texte.
- Moyen privilégié de protéger et d'évaluer l'authenticité des documents sur le long terme, le processus d'évaluation doit être précisément documenté. Dans la mesure où le processus de conservation débute avec la production des documents, la responsabilité de cette documentation revient tant au producteur qu'au service en charge de leur conservation. Par le passé, la stabilité

¹³ Voir Kenneth Thibodeau et al., « Part Three – Trusting to Time : Preserving Authentic Records in the Long Term : Preservation Task Force Report », *ibid.*, 99-116. Une version électronique est disponible à l'adresse suivante : http://www.interpares.org/book/interpares_book_f_part3.pdf.

de la plupart des formes de documents analogiques permettait souvent au producteur de se désintéresser du sort de ses documents jusqu'au versement des archives définitives au service ou à l'institution en charge de leur conservation.

Lorsque les chercheurs d'InterPARES 2 ont élaboré les différentes méthodes de recherche qui seraient utilisées pendant le projet, l'idée était de déterminer si les études de cas, les travaux de modélisation ou les enquêtes d'InterPARES 2 révéleraient des théories ou des pratiques en décalage avec les recommandations d'InterPARES 1. Dans l'ensemble, les conclusions d'InterPARES 1, établies à partir de l'étude de grandes bases de données et d'applications de gestion des documents, n'étaient pas en rupture avec les connaissances et les pratiques archivistiques existantes. Il est apparu que ces bases de données et ces applications produisaient, pour l'essentiel, des représentations numériques de types de documents établis de longue date dans les environnements d'archivage papier, comme les dossiers et les documents textuels. L'étude de formes plus novatrices de documents numériques par les chercheurs d'InterPARES 2 suggérerait-elle la nécessité de nouveaux critères d'évaluation, ou l'ajout d'étapes aux pratiques d'évaluation actuelles, ou encore l'élaboration de nouvelles politiques ou pratiques de conservation ?

L'équipe de recherche

On trouvera ci-dessous la liste des chercheurs et des assistants de recherche ayant participé à l'équipe thématique 3 au cours du projet¹⁴ :

Présidents et co-présidents :

Yvette Hackett et Sally Hubbard, Juin 2005 – Déc. 2006 (Co-présidentes)

Hans Hofman et Sally Hubbard, Fév. 2004 - Juin 2005 (Co-présidents)

Kenneth Thibodeau, Janvier 2002 – Déc. 2003 (Président)

Chercheurs :

Howard Besser, New York University, États-Unis — Groupe de travail 3.1

Ann Butler, New York University, États-Unis — Groupe de travail 3.1

Kevin Glick, Yale University, États-Unis — Groupe de travail 3.2

Elaine Goh, National Archives of Singapore — Groupe de travail 3.3

¹⁴ La composition de l'équipe a beaucoup changé au cours des cinq ans du projet. Signalons en particulier que les chercheurs suivants, intéressés par les travaux de l'équipe thématique 3, n'ont pu y prendre part sur toute la durée du projet : Filip Boudrez, Archives municipales d'Anvers/the DAVID Project, Belgique ; Michèle Cloonan, Simmons College, États-Unis ; Margaret Hutchison, British Columbia Archives, Canada ; Glenn Isaac, British Columbia Archives, Canada ; Rick Kopak, The University of British Columbia, Canada ; Rich Lysakowski, CENSA, États-Unis ; Jean-Stéphane Piché, Library and Archives Canada ; Shelby Sanett, Amigos Library Services, Inc., États-Unis ; Lynne Tibbitt, British Columbia Archives, Canada ; et Bill Underwood, Georgia Tech Research Institute, États-Unis.

Yvette Hackett, Library and Archives Canada — Groupe de travail 3.1
 Babak Hamidzadeh, Library of Congress, États-Unis — Groupe de travail 3.2
 P.C. Hariharan, Systems Engineering & Security, Inc., États-Unis — Groupe de travail 3.2
 Ken Hawkins, National Archives and Records Administration, États-Unis — Groupe de travail 3.3
 Hans Hofman, Archives nationales des Pays-Bas — Groupe de travail 3.3
 Sally Hubbard, Getty Institute, États-Unis — Groupe de travail 3.1
 Mary Ide, WGBH, États-Unis — Groupe de travail 3.1
 Randal Luckow, Turner Broadcasting, États-Unis — Groupe de travail 3.1
 Richard Marciano, San Diego Supercomputer Center, États-Unis — Groupe de travail 3.3
 Evelyn McLellan, Insurance Corporation of British Columbia, Canada — Groupe de travail 3.3
 Reagan Moore, San Diego Supercomputer Center, États-Unis — Groupe de travail 3.2
 Isabella Orefice, Associazione Nazionale Archivistica Italiana — Groupe de travail 3.1
 Jim Suderman, Archives of Ontario, Canada — Groupe de travail 3.3
 Kenneth Thibodeau, National Archives and Records Administration, États-Unis — Groupe de travail 3.3
 James Turner, Université de Montréal, Canada — Groupe de travail 3.1

Assistants de recherche :

Tom Anderson, The University of British Columbia, Canada
 Patsy Baudoin, Simmons College, États-Unis
 Carolyn Casenas, The University of British Columbia, Canada
 Natalie Catto, The University of British Columbia, Canada
 Alan Doyle, The University of British Columbia, Canada
 Adam Farrell, The University of British Columbia, Canada
 Fiorella Foscarini, The University of British Columbia, Canada
 Peggy Heger, The University of British Columbia, Canada
 Tracey Krause, The University of British Columbia, Canada
 Karen Langle, The University of British Columbia, Canada
 Catherine Miller, The University of British Columbia, Canada
 Luke Meagher, The University of British Columbia, Canada
 Jennifer Meehan, The University of British Columbia, Canada
 Shaunna Moore, The University of British Columbia, Canada
 Elisheba Muturi, The University of British Columbia, Canada

Carolyn Petrie, The University of British Columbia, Canada

Brian Trembath, The University of British Columbia, Canada

Sherry Xie, The University of British Columbia, Canada

Questionnaire de recherche

Les quatre questions de recherche auxquelles il était prévu que réponde le thème 3 étaient définies comme suit dans la proposition de recherche initiale :

1. Les concepts, méthodes et modèles d'évaluation élaborés par InterPARES 1 pour les documents administratifs et juridiques produits dans des bases de données et des systèmes de gestion des documents s'appliquent-ils à l'évaluation des documents d'activités artistiques, scientifiques et administratives résultant de l'utilisation des technologies examinées par InterPARES 2 ?
2. Les concepts, méthodes et modèles de conservation élaborés par InterPARES 1 pour les documents administratifs et juridiques produits dans des bases de données et des systèmes de gestion des documents s'appliquent-ils à la conservation des documents d'activités artistiques, scientifiques et administratives résultant de l'utilisation des technologies examinées par InterPARES 2 ?
3. Quels paradigmes de conservation peuvent être appliqués à toutes les activités et toutes les technologies ? Quels paradigmes de conservation sont requis pour les types particuliers de documents résultant de chaque activité ?
4. Quelles métadonnées sont nécessaires pour soutenir l'évaluation et la conservation de documents numériques authentiques résultant de chaque activité ?

Comme cela a déjà été indiqué, les analyses envisagées pour la question 1 et la question 2 ont été transférées à l'équipe transversale Modélisation. Toutefois, il demeurerait dans les attributions des chercheurs de l'équipe thématique 3 d'identifier toute modification des activités d'évaluation et/ou de conservation qui pourrait être nécessaire dans les environnements explorés par les études de cas d'InterPARES 2, par rapport aux environnements bureaucratiques plus traditionnels ciblés par InterPARES 1. Il convenait en particulier de déterminer si les procédures d'évaluation ou de conservation devraient faire l'objet d'ajustements ou d'évolutions afin de prendre en charge les technologies numériques, par-delà ceux déjà identifiés par InterPARES 1.

Le changement de paradigme observé lors des recherches d'InterPARES 1 s'est traduit par une évolution du vocabulaire de la conservation numérique, le concept de conservation du « document d'archives » cédant la place à la reconnaissance de l'importance de l'identification et de la conservation des « composants numériques » et de leurs relations. Il était particulièrement important de déterminer si l'utilisation de systèmes interactifs, expérientiels ou dynamiques exigerait une évolution du même ordre de la façon dont les archivistes envisagent leur métier.

Enfin, on a jugé que la réponse à la quatrième question découlerait naturellement de l'examen des études de cas par les archivistes. Si ces derniers se trouvaient dans l'impossibilité d'évaluer les documents des études de cas ou de définir une politique de conservation adaptée, cela tendrait à indiquer que les métadonnées étaient insuffisantes et, partant, qu'il convenait de définir le type de métadonnées requises.

[...]

Conclusions

L'analyse des études de cas et des études générales indique que les producteurs sont encore trop nombreux à négliger la conservation sur le long terme de leurs fichiers numériques, qu'il s'agisse de documents statiques, dynamiques, probants, expérientiels, interactifs ou patrimoniaux. Du fait de la diversité et de l'ampleur des études de cas et des études générales, tous les types d'approches et d'attitudes ou presque ont été observés au moins une fois. Cependant, on peut établir certaines distinctions entre d'une part les particuliers et les petites organisations, et d'autre part les grandes entreprises et les organismes et services publics. Les études de cas d'InterPARES 2 conduites dans la sphère de l'administration ont débouché sur des conclusions similaires à celles d'InterPARES 1, révélant chez les acteurs du domaine une sensibilisation et une connaissance des bonnes pratiques de production et de maintenance des documents d'archives, papier ou électroniques, bien supérieures à celles observées dans les autres domaines. À la différence des particuliers et des petites organisations, les grandes institutions disposent aussi généralement de systèmes d'archivage papier, vers lesquels elles peuvent se tourner pour identifier l'exemplaire papier faisant autorité et l'intégrer au système de gestion des documents analogiques. Les sites web, toutefois, échappent à cette pratique ; petites et grandes organisations les traitent souvent comme des objets éphémères, en ce sens qu'aucune mesure ou presque n'est prise pour en conserver des versions fixes ; cela étant, un certain nombre

d'organisations considèrent les sites web soit comme un « système d'archivage¹⁵ » soit comme un site patrimonial où préserver leur travail¹⁶.

En ce qui concerne les individus et les petites organisations, les pratiques de conservation sont inexistantes ou presque, de même que toute politique de maintenance, ne serait-ce qu'élémentaire¹⁷. Dans de nombreux cas, lorsque des procédures de maintenance existaient, elles étaient davantage le fruit d'un heureux hasard que d'un acte délibéré de conservation. La politique la plus utilisée est à n'en pas douter celle de la « redondance », qui consiste à faire une copie du document ailleurs que sur l'ordinateur du producteur ou le disque dur du serveur. Dans les grandes organisations, les sauvegardes régulières réalisées par le personnel informatique étaient et sont désormais pratique courante¹⁸. Quant aux petites organisations et aux individus, une forme ou une autre de sauvegarde était souvent mentionnée¹⁹. La disparition de l'ancien concept DOS de formats de « sauvegarde », qui étaient très sensibles aux mises à jour du système d'exploitation, constitue à cet égard un progrès indéniable. Les coûts de stockage diminuant, davantage de copies sont des copies de base, c'est-à-dire des copies sauvegardées dans le format de fichier dans lequel les documents ont été produits à l'origine, ou dans lequel ils ont été utilisés et sauvegardés pour la dernière fois, les rendant ainsi (théoriquement) plus facilement lisibles par l'être humain dans l'environnement de travail habituel du producteur. La structure des répertoires et les noms des fichiers sont immédiatement accessibles sans qu'il soit

¹⁵ Voir par exemple : Heny Daniel et Cara Payne (2004), « InterPARES 2 Project – Case Study 02 Final Report : Performance Artist Stelarc », 7-13, http://www.interpares.org/display_file.cfm?doc=cs02_final_report.pdf.

¹⁶ Voir par exemple : Martine Cardin (2004), « InterPARES 2 Project – Case Study 01 Final Report : Arbo Cyber, théâtre, (?) », http://www.interpares.org/display_file.cfm?doc=ip2_cs01_final_report_english.pdf. Le texte français original est disponible à l'adresse suivante : http://www.interpares.org/display_file.cfm?doc=ip2_cs01_final_report_french.pdf.

¹⁷ Huit politiques élémentaires de maintenance sont décrites en annexe aux *Principes directeurs à l'usage des producteurs* élaborés par les chercheurs du thème 3 et présentées en annexe du présent rapport (voir l'Annexe 21c). Ces politiques sont les suivantes : (1) définition des responsabilités ; (2) mise en place d'une infrastructure technique adaptée; (3) maintenance, support et remplacement des systèmes ; (4) migration périodique des données sur de nouveaux supports de stockage ; (5) respect de conditions climatiques appropriées pour les supports de stockage ; (6) redondance et localisation géographique ; (7) sécurité des systèmes ; et (8) plan de prévention des sinistres et plan de restauration.

¹⁸ Voir, par exemple, pour le domaine artistique, James Turner et al. (2004), « InterPARES 2 Project – Case Study 09(3) Final Report : Digital Moving Images – Commercial Film Studio », 8-10, http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_final_report.pdf ; pour le domaine scientifique, Tracey P. Lauriault et Yvette Hackett (2005). « InterPARES 2 Project – Case Study 06 Final Report : Cybercartographic Atlas of Antarctica », 17, http://www.interpares.org/display_file.cfm?doc=ip2_cs06_final_report.zip ; et pour le domaine administratif, Elaine Goh (2005). « InterPARES 2 Project – Case Study 21 Final Report : The Electronic Filing System (EFS) of the Supreme Court of Singapore », 32, http://www.interpares.org/display_file.cfm?doc=ip2_cs21_final_report.pdf.

¹⁹ Voir, par exemple, pour les sauvegardes : (1) sur CD-ROM ou DVD dans Daniel et Payne, « Case Study 02 Final Report », op. cit., 13, et Bart Ballaux (2005), « InterPARES 2 Project – Case Study 26 Final Report : MOST Satellite Mission – Preservation of Space Telescope Data », 7, http://www.interpares.org/display_file.cfm?doc=ip2_cs26_final_report.pdf ; et (2) sur un autre ordinateur dans Sydney Fels et Seth Dalby (2004), « InterPARES 2 Project – Case Study 15 Final Report : *Waking Dream* », 4, http://www.interpares.org/display_file.cfm?doc=ip2_cs15_final_report.pdf. Voir également Ballaux, « Case Study 26 Final Report », op. cit.

nécessaire de décompresser pour rouvrir la copie de sauvegarde ou de protection, souvent compressée. Dans le cas de fichiers informatiques très lourds, cependant, il n'est pas exclu que la compression interdise cette simplification.

Ces copies redondantes étaient également, dans certains cas, stockées sur des sites distants. L'exemple le plus complet de cette démarche est l'étude de cas consacrée au Système de dépôt électronique (EFS) de la Cour suprême de Singapour, dont le rapport indique :

Il y a deux sauvegardes des données, l'une quotidienne, l'autre hebdomadaire. Elles sont conservées sur un site distant afin de garantir la restauration complète des données en cas de défaillance du système. Par exemple, chaque fois qu'un document est versé par un cabinet juridique, les documents sont stockés sur trois disques. Le premier disque réside de façon permanente dans le juke-box pour faciliter l'accès en ligne aux informations. Le deuxième disque est transféré sur un site de stockage distant en fin de semaine et le troisième disque est déplacé sur un site distant lorsqu'il est plein²⁰.

Dans certains cas, les copies redondantes étaient conservées au domicile du producteur et à son bureau²¹, ni l'un ni l'autre lieu n'offrant des conditions climatiques – température, hydrométrie, pollution de l'air – stables. Qui plus est, bien que ce type de politique « informelle » de stockage réparti offre effectivement une protection accrue contre le vandalisme, le vol ou l'incendie, elle se révélerait sans doute de peu d'utilité en cas de catastrophe plus importante, tremblement de terre ou inondation par exemple²². Hormis la récupération des données offerte par ces procédures rudimentaires de *sauvegardes sélectives* de données ou de fichiers²³, il n'existait quasiment aucun plan de prévention et

²⁰ Goh, « Case Study 21 Final Report », op. cit. Voir aussi Ballaux, « Case Study 26 Final Report », op. cit., pour un exemple de l'utilisation des procédures de sauvegarde sur site distant par un producteur dans la sphère scientifique.

²¹ Voir par exemple la politique du producteur dans l'étude de cas n° 14 : des copies des données GIS et des fichiers associés sont « maintenues » de manière *ad hoc*, « sur l'ordinateur personnel du producteur, le serveur de l'entreprise et/ou copiés sur des CD-ROM » (Richard Pearce-Moses, Erin O'Meara et Randy Preston (2004), « InterPARES 2 Project – Case Study 14 Final Report : Archaeological Records in a Geographical Information System : Research in the American Southwest », 28, http://www.interpares.org/display_file.cfm?doc=ip2_cs14_final_report.pdf).

²² On pourrait dire la même chose d'accords de stockage répartis plus officiels, comme celui mentionné dans l'étude de cas n° 21, selon la nature des installations de stockage distant utilisées et leur proximité géographique par rapport au site.

²³ Par opposition aux *sauvegardes systèmes complètes*, où le système d'exploitation, tous les logiciels et toutes les données et tous les fichiers sont sauvegardés de sorte que l'intégralité du système peut être restaurée en cas de défaillance massive du système (voir *Principes directeurs à l'usage des producteurs*, principe 8.A, Annexe 20). Deux exceptions notables au recours exclusif à des stratégies de sauvegarde sélective ont été observées dans les études de cas ; une dans la sphère scientifique, l'autre dans la sphère de l'administration. Dans l'étude de cas n° 26 (MOST Satellite Mission), on peut lire « dans la mesure où les chercheurs de MOST travaillent presque exclusivement avec des logiciels spécifiques, la conservation du logiciel est une question centrale. Des sauvegardes des différents programmes utilisés par le projet sont faites. En outre, si quoi que ce soit est modifié (c'est-à-dire, ajouté) dans un programme, l'ancienne version du logiciel est toujours conservée. De cette façon, les chercheurs peuvent toujours recréer les résultats créés antérieurement. » (Ballaux, « Case Study 26 Final Report », op. cit., 7-8). Dans l'étude de cas n° 12 (Base de données du Traité de l'Antarctique), il est indiqué que « les objets des documents numériques de la base de données sont maintenus sur des serveurs avec des copies de sauvegarde sur d'autres disques durs ainsi que sur des WebCDservers... qui

de gestion des catastrophes chez les producteurs²⁴.

Certaines études de cas et études générales ont également confirmé ce que les chercheurs du thème 3 savaient déjà par expérience. La plupart des producteurs ne vont pas au-delà d'une maîtrise et d'une utilisation rudimentaires des technologies qu'ils doivent utiliser. Peu d'utilisateurs deviennent experts ou apprennent à évaluer les implications des choix de configuration, des additifs développés par des tierces parties, ou des fonctionnalités pertinentes dans l'environnement informatique qu'ils ont choisi²⁵ pour la conservation des documents sur le long terme.

Les chercheurs du thème 3 ont noté que, dans le domaine artistique tout particulièrement, le matériel, et notamment les périphériques, peut jouer un rôle essentiel dans la reproduction exacte et authentique d'une œuvre d'art²⁶ par exemple. Toutefois, cette dépendance vis-à-vis d'un matériel informatique donné peut empêcher toute conservation, en particulier si la responsabilité de la conservation est transférée à un archiviste tiers de confiance qui n'a pas accès au matériel requis et/ou pas la capacité ou les ressources pour le maintenir. Les chercheurs du thème 2 ont donc recommandé que les producteurs s'efforcent dans la mesure du possible de « éliminer toute dépendance à l'égard du matériel informatique en transférant les fonctionnalités de celui-ci à un logiciel (c'est-à-dire en utilisant une application logicielle qui simule les actions d'une partie du matériel » [car] « ceci constitue une manière plus stable de conserver les fonctions lorsque les matériels deviennent obsolètes. »²⁷

reproduisent l'intégralité des fonctionnalités et du contenu du site Web... Le WebCDserver a été utilisé une fois pour restaurer le site de la Base de données du Traité sur l'Antarctique. » (Paul Arthur Berkman et al. (2005), « InterPARES 2 Project – Case Study 12 Final Report : Antarctic Treaty Searchable Database », 39. Voir :

http://www.interpares.org/display_file.cfm?doc=ip2_cs12_final_report.pdf). Bien qu'il ne soit fait mention ci-dessus que de la sauvegarde des « objets des documents numériques », il semble que des copies du système d'exploitation et de l'application logicielle Digital Integration System soient également incluses dans la sauvegarde sur WebCDservers, dont le producteur dit plus loin qu'il contient « des copies parfaitement exécutables des sites web » (Ibid., 18).

²⁴ L'étude de cas n° 9(3) (Studio films publicitaires) est la seule à mentionner un plan de prévention et de gestion des catastrophes (voir Turner et al., « Case Study 09(3) Final Report », op. cit., 15 (voir la réponse à la question 3.2).

²⁵ C'est ce qu'illustre par exemple l'étude générale n° 9 (Pratiques d'archivage numérique des archéologues utilisant un SIG), où il est indiqué que « plusieurs participants [à l'enquête] ont reconnu se sentir un peu écrasés par le potentiel analytique de leur SIG du fait, essentiellement, du manque de formation au SIG. » Comme le souligne le rapport, « si c'est vrai, cela pourrait avoir des répercussions pour la question plus large de la conservation sur le long terme des documents des SIG. De fait, il est possible que plus la personne se sente aliénée, déconnectée et/ou intimidée par les projets de SIG, moins elle éprouve le sentiment d'obligation, d'initiative et/ou de compétence nécessaire pour aborder efficacement les exigences de conservation sur le long terme du projet. » (Preston, « General Study 09 Final Report », op. cit., 32).

²⁶ Voir dans l'étude de cas n° 13 (*Obsessed Again...*) le compte rendu de la tentative de chercheurs du thème 2 de ressusciter une œuvre électroacoustique victime d'obsolescence technologique. On trouvera également un résumé de cette expérience dans le Rapport de l'équipe thématique 2, section « A strategy for preventing technological obsolescence of an artistic work », disponible à l'adresse suivante : http://www.interpares.org/display_file.cfm?doc=ip2_book_part_3_domain2_task_force.pdf.

²⁷ Voir *Principes directeurs à l'usage des producteurs*, principe 9, Annexe 20.

Les chercheurs ont également noté que plusieurs producteurs se désintéressent purement et simplement de la conservation de leurs documents, bien que pour diverses raisons. Dans certains cas, les producteurs (en particulier dans la sphère artistique mais également dans la sphère scientifique) ont indiqué qu'il serait plus facile et moins coûteux de re-créeer plus tard certains types de contenus sous forme numérique que d'essayer de garder et de convertir ou de fournir une plateforme d'émulation pour des documents très spécialisés. C'est ce qu'illustre le passage ci-dessous, tiré de l'étude de cas n° 9(3) (Studio de films publicitaires) du domaine artistique :

Dans la mesure où des changements importants de matériel et de logiciels sont généralement intervenus depuis que l'œuvre d'art a été créée, il est jugé plus économique dans l'environnement étudié de re-créeer l'œuvre si on doit la réutiliser, que de se lancer dans une démarche de migration ou de mise à niveau pour le cas où on serait amené à la réutiliser. Cette approche permet de limiter le temps, l'énergie et les moyens financiers consacrés à la conservation des objets numériques sur le long terme et d'éviter d'avoir à mettre en place des politiques de conservation qui répondent aux problèmes d'obsolescence du matériel et des logiciels²⁸.

Ce point de vue fait écho à celui du producteur de l'étude de cas scientifique n° 14 (Documents archéologiques dans des systèmes d'information géographique) qui, en réponse à une question sur les fichiers intermédiaires (par exemple, les algorithmes de coût surface) qu'il crée et leurs relations avec ses travaux sur le SIG, a indiqué :

Je n'accorde sans doute pas beaucoup de valeur à la plupart de ces documents intermédiaires, parce qu'ils sont faciles à recréer... Chaque fois que j'essaie d'apprendre à quelqu'un les rudiments du SIG, les gens me disent, voilà les données dont je dispose et voilà ce que je veux obtenir. Et la première [chose] que je leur [dis] toujours est que s'ils posent la question à dix personnes qui travaillent sur le SIG, ils obtiendront dix réponses différentes. Et comme chacune de ces façons de procéder produira des fichiers intermédiaires différents, il me semble que pour la plupart ils n'ont guère de valeur. Comme je l'ai dit, ils sont faciles à reproduire. Si je le veux, je peux toujours faire un autre fichier de pente, si j'en perdais un ou s'il y avait un problème²⁹.

Dans d'autres cas, certains artistes ne souhaitent pas que leur travail soit conservé sur le long terme si cela risque de compromettre certaines caractéristiques qu'ils considèrent comme fondamentales pour l'essence de leurs œuvres, ainsi du caractère éphémère et mouvant d'œuvres qui changent chaque fois qu'elles sont

²⁸ Turner et al., « Case Study 09 (3) Final Report », op. cit., 4.

²⁹ Pearce-Moses et al., « Case Study 14 Final Report », op. cit., 31.

reproduites. C'est ce qu'illustre l'étude de cas n° 15 *Waking Dream*, consacrée à une performance multimédia. Une des interprètes et des créatrices de l'œuvre, Takahashi Sachiyo, « considère *Waking Dream* comme une œuvre de performance définie par son rôle en tant que performeuse », conception qui, toujours selon elle, « exclut toute performance de l'œuvre sans elle [comme performeuse]³⁰. » Selon cette interprétation, il serait impossible de « conserver » la capacité de reproduire l'œuvre ou, à tout le moins, une reproduction authentique de l'œuvre, après la mort du performeur.

Ces observations ont conduit les chercheurs d'InterPARES à continuer à explorer et affiner la distinction entre les documents numériques qui sont des œuvres ou *documentent* des représentations, et les objets numériques qui *rendent possible* la reproduction ultérieure d'une œuvre ou d'une performance – concept proche du « négatif » en photographie – ou fournissent des *instructions* concernant l'exécution d'une œuvre ou d'une performance – concept proche de la « partition » en musique –, à quoi il faut ajouter que ces *documents instrumentaux* et *d'instructions* joueront réellement un rôle actif ou instructif dans la re-présentation de l'œuvre ou de la performance. À la différence des documents d'instructions, qui « sont destinés à être lus par l'être humain et sont donc matérialisés en étant reproduits dans une forme lisible par l'être humain à partir de composants numériques stockés³¹ », une caractéristique clé des documents instrumentaux, que ce soit dans le domaine artistique³² ou le secteur manufacturier où d'autres exemples ont été trouvés³³, est qu'ils contiennent des instructions destinées à être exploitées par une machine et, par conséquent,

accomplissent leur finalité sous la forme numérique dans laquelle ils ont été stockés et, inversement, ne peuvent atteindre cet effet s'ils sont convertis dans un format lisible par l'être humain. En outre,

³⁰ Fels et Dalby, « Case Study 15 Final Report », op. cit., 8.

³¹ Luciana Duranti et Kenneth Thibodeau (2006). « The Concept of Record in Interactive, Experiential and Dynamic Environments : the View of InterPARES », *Archival Science* 6(1) : 60. Pour la traduction française de cet article, voir l'annexe 2.

³² Par exemple, le programme créé et utilisé pour contrôler le dispositif télécommandé utilisé par un performer dans *Waking Dream* (voir Fels et Dalby, « Case Study 15 Final Report », 2 (contexte technologique)) et les différents patches logiciels spécifiquement écrits pour *Obsessed Again*, et notamment un patch synthétiseur qui contrôle les sons que l'ordinateur fait jouer au synthétiseur en réponse aux notes jouées par le basson (voir J. Scott Amort (2004). « InterPARES 2 Project – Case Study 13 Final Report : *Obsessed Again...* », 3 (voir la réponse à la question 4)). Disponible à l'adresse suivante : http://www.interpares.org/display_file.cfm?doc=ip2_cs13_final_report.pdf.

³³ Par exemple, le fichier système CAD Pro-Engineer utilisé pour créer les fichiers CAD originaux dans l'étude de cas n° 19 (voir Kenneth Hawkins (2006). « InterPARES 2 Project – Case Study 19 Final Report : Preservation and Authentication of Electronic Engineering and Manufacturing Records », 6. Voir : http://www.interpares.org/display_file.cfm?doc=ip2_cs19_final_report.pdf.) répond à la définition du document instrumental en ce sens qu'il constitue un ensemble fixe d'instructions exploitables seulement par une machine, qui pilote des actions ultérieures et fournit des instructions à des robots d'ingénierie, etc. L'ajout de métadonnées sémantiques ou opérationnelles aux itérations ultérieures des mêmes informations fixes permet d'autres actions ; à savoir, « l'interrogation » des formats à connaissances augmentées par des moteurs d'inférence.

tant qu'ils sont actifs, les documents instrumentaux doivent être maintenus dans les systèmes où ils ont été produits – ou dans des systèmes dotés de fonctionnalités identiques. Autrement, ils ne produiraient pas ou ne faciliteraient pas les interactions, expériences, performances ou autres processus qu'ils étaient destinés à générer³⁴.

D'autres producteurs encore estiment qu'ils peuvent attendre pour prendre des mesures de conservation – ou s'en dispenser purement et simplement – parce qu'ils sont convaincus que l'industrie informatique sortira de son chapeau la solution requise, lorsqu'ils en auront besoin. « Faire confiance au fournisseur » est une stratégie qui peut fonctionner lorsqu'une technologie a été largement adoptée et qu'il existe sur le marché un produit équivalent vers lequel les clients pourraient se tourner. N'oublions pas pour autant que d'innombrables supports ont été abandonnés par les fabricants depuis les cartouches 8 pistes, quand bien même des milliers de consommateurs se voyaient contraints de migrer toutes leurs collections de musique ou de vidéo sur une nouvelle technologie.

Cette confiance en son prestataire, ou en l'industrie informatique en général, semble également être ce qui motive l'adoption de la technologie de la signature numérique en dépit des interrogations qui pèsent sur la pérennité de cette solution d'authenticité³⁵. C'est d'ailleurs ce dont semble témoigner l'étude de cas consacrée à l'informatisation du livre foncier d'Alsace-Moselle. L'entité administrative dédiée chargée de superviser l'informatisation (et la numérisation) du livre foncier,

a l'obligation légale d'assurer la communication dans le temps des données du livre foncier et de garantir leur valeur probante, conformément à l'article 1316-1 du Code civil, quelles que soient les évolutions technologiques, [pourtant] [elle] n'a pas envisagé le problème de la maintenance des

³⁴ Duranti et Thibodeau, « Le concept de document d'archives », op. cit. Comme Duranti et Thibodeau le précisent, « on peut citer comme exemples de documents instrumentaux les correctifs logiciels qui permettent à un instrument de musique d'interagir avec un ordinateur, les logiciels des sites de marketing qui interprètent les données relatives aux actions du visiteur sur le site afin de déterminer quels éléments de contenu devront lui être présentés lors de sa prochaine visite, et les agents logiciels qui permettent à des applications métier en interaction d'exécuter des transactions de manière indépendante. » (Ibid., 59).

³⁵ Comme le soulignent les *Principes directeurs à l'usage des producteurs*, principe 6 (voir l'annexe 20), la signature numérique est une technique d'authentification dépendante de la technologie et, à ce titre, n'est certes pas à l'abri de l'obsolescence. Par nature, les signatures électroniques ne peuvent en effet, à l'heure actuelle, être migrées vers de nouvelles applications logicielles ou des applications mises à niveau, avec les documents auxquels elles sont jointes. Qui plus est, la durée de vie des signatures numériques et des autres technologies d'authentification est dans certains cas beaucoup plus courte que la durée de maintenance requise pour un document même temporaire ne nécessitant pas de migration, parce que la technologie d'authentification évolue rapidement. Plusieurs politiques de conservation (voir l'annexe 21c, partie B) exigeraient à plus ou moins longue échéance la conversion dans un nouveau format d'un document comportant une signature numérique. Il est peu probable que le document et sa signature puissent être revalidés par l'infrastructure à clé publique puisque l'empreinte (somme de contrôle) aura changé. En conséquence, les producteurs et les services en charge de la conservation sur le long terme sont encouragés à utiliser dans la mesure du possible des technologies d'authentification indépendantes de la technologie, s'appuyant sur des politiques et des procédures administratives efficaces pour la production, la maintenance et la conservation des documents d'archives.

objets numériques, sauf via les mécanismes proposés par les fournisseurs du système au moment de la mise à niveau du système de gestion de la base de données. En outre, n'a pas été envisagé non plus le problème de la maintenance de la valeur probante des signatures numériques au fil de l'évolution technologique³⁶.

D'autres producteurs encore se disent capables de « surveiller eux-mêmes » l'évolution des technologies et des normes influant sur la conservation des documents d'archives sur le long terme et d'adapter leurs systèmes et leurs pratiques de conservation en conséquence, au rythme des évolutions technologiques. L'étude de cas sur le Système de dépôt en ligne de la Cour suprême de Singapour en offre une excellente illustration. Le producteur (la Cour suprême de Singapour) décrit sa démarche actuelle de conservation comme une tentative d' « anticiper les nouvelles normes/technologies avant que la technologie en place ne devienne totalement obsolète », afin de se donner « suffisamment de temps pour migrer vers les nouvelles normes/technologies sans perdre de données³⁷. » Bien que la veille technologique et l'adoption de mesures contre l'obsolescence du matériel et des logiciels soient au nombre des recommandations clés des *Principes directeurs à l'usage des producteurs*³⁸, elles ne seront efficaces que si elles s'inscrivent dans un cadre intégré de politiques et de procédures. Malheureusement, il ne semble pas que ce soit le cas de la Cour suprême de Singapour, le rapport indiquant un peu plus loin que bien que « la cour reconnaisse la nécessité de se préoccuper de la conservation sur le long terme des documents d'archives électroniques... aucune politique n'est actuellement en place *parce que la cour considère que le système est très répandu*³⁹. » La raison invoquée ici par le producteur pour différer l'élaboration d'une politique de conservation sur le long terme justifie l'inquiétude des chercheurs du thème 3 quant aux pratiques de conservation d'un grand nombre des producteurs étudiés. Cette attitude insouciant dénote une incompréhension fondamentale de la dimension temporelle du problème et du fait qu'une conservation sur le long terme efficace débute au moment de la production du document.

Les chercheurs du thème 3 ont observé que les technologies adoptées par les innovateurs, indépendamment de leur domaine d'activité, étaient propriétaires⁴⁰ et souvent spécifiques. Dans de

³⁶ Jean-François Blanchette, Françoise Banat-Berger et Geneviève Shepherd (2004). « InterPARES Project – Case Study 18 Final Report : Computerization of Alsace-Moselle's Land Registry », 27. Voir http://www.interpares.org/display_file.cfm?doc=ip2_cs18_final_report.pdf. Voir également, pour la version en français : <http://www.cairn.info/revue-document-numerique-2004-2-page-63.htm>.

³⁷ Goh, « Case Study 21 Final Report », op. cit., 33.

³⁸ Voir Annexe 20, principe 9.

³⁹ Goh, « Case Study 21 Final Report », op. cit., 38 (c'est nous qui soulignons).

⁴⁰ Propriétaire, soit dans le sens de logiciels dont le code source est la propriété d'un éditeur privé, soit dans le sens de logiciels spécifiques, dont le code source est développé « en interne » et n'est pas librement disponible. Le logiciel dédié

nombreux cas, le but du travail de ce type de producteurs est d'explorer, de tester et de repousser les limites des technologies existantes, qu'il s'agisse de matériel ou de logiciels. Dès lors, il ne saurait être question d'attendre l'adoption de normes internationales ou la diffusion de formats⁴¹ « ouverts ». Ces environnements nécessitent une grande interopérabilité⁴², à tout le moins jusqu'à ce que l'adoption à grande échelle de la technologie génère des normes, à supposer que cela arrive un jour. L'équipe transversale Stratégie et orientations indiquait dans son rapport final que « les mesures prises pour assurer l'interopérabilité des systèmes remplissent bon nombre des finalités de la conservation⁴³. »

Beaucoup de scientifiques considèrent la publication d'un article dans une revue comme un moyen de conservation, plutôt que d'essayer de préserver les antécédents numériques ayant conduit à l'article publié. Ce point de vue a été exprimé par beaucoup de participants de l'étude générale sur les pratiques d'archivage électronique des archéologues utilisant un SIG lorsqu'il leur a été demandé quels étaient selon eux les éléments ou les résultats les plus importants de leurs projets de SIG à conserver à des fins d'utilisation ou de consultation ultérieures, et pourquoi. Comme l'indique le rapport final de l'étude, « pour de nombreux participants, la seule publication constitue une conservation sur le long terme suffisante de leurs travaux⁴⁴. » À la lumière de ce constat, le rapport tire un peu plus loin la sonnette d'alarme :

En se concentrant sur la conservation des rapports finaux (et des types associés de documents, comme les articles publiés), au détriment d'une documentation plus complète des projets en cours, les archéologues courent le risque de ne pas conserver les types de documents qui permettront que les objectifs clés de la conservation, identifiés par les participants à l'enquête eux-mêmes, - à savoir, réutilisation, vérification, reproduction et justification des projets/données – soient effectivement atteints⁴⁵.

Cette attitude n'est pas étrangère aux employés de l'administration qui privilégient parfois la conservation d'une étude, au détriment des données et recherches l'ayant nourrie. C'est notoirement insuffisant, dans la mesure où la perte de données dans une forme manipulable interdit de vérifier les

utilisé pour traiter les données satellites dans l'étude de cas n° 26 (MOST Satellite Mission) en est un exemple.

⁴¹ Pour une discussion approfondie sur l'utilisation et la définition du concept de « format de fichier » dans le contexte des institutions archivistiques, et notamment les termes « format de fichier standard » et « format de fichier ouvert » (également appelés formats « open source » et « non propriétaires »), voir McLellan, « General Study 11 Final Report », op. cit.

⁴² La base de données terminologique d'InterPARES 2 définit l'interopérabilité comme « la faculté d'une application/d'un système de communiquer ou travailler avec une/un autre » (http://www.interpares.org/ip2/ip2_terminology_db.cfm).

⁴³ Rapport de l'équipe transversale Stratégie et orientations, 319.

⁴⁴ Preston, « General Study 09 Final Report », op. cit., 4.

⁴⁵ Ibid., 91.

résultats, est un obstacle à la duplication éventuelle de l'expérience de collecte des données et ampute les séries chronologiques d'une partie de leurs données. En outre, l'argument « publication » a perdu une grande part de sa pertinence avec le passage aux revues et aux rapports électroniques, puisqu'il est clair que la publication, même sous forme numérique, ne peut se substituer aux données originales. La nouvelle plate-forme numérique pour la conservation des revues électroniques a créé un environnement où les données associées peuvent tout à fait être préservées, si on le souhaite, et être rattachées à la publication.

Plusieurs entretiens conduits lors des études de cas montrent que, pour de nombreux producteurs d'objets numériques, « les mettre sur le web » est un moyen d'assurer leur survie sur le long terme – comme on plaçait hier sa confiance dans la publication papier. Un point de vue pour le moins sujet à caution quand on sait qu'aucune organisation n'essaye même de conserver l'intégralité du web et les millions de changements apportés toutes les secondes à son contenu⁴⁶. Ces attitudes reflètent la conviction la plus répandue mise en lumière par les études de cas : la conservation numérique est « l'affaire de quelqu'un d'autre⁴⁷. » Ce « quelqu'un d'autre » désignant aussi bien un individu donné – quelqu'un d'autre dans l'entreprise en est chargé – que quelqu'un en général – par exemple, une institution publique ou une autre organisation ou entité de recherche et ayant des ressources, s'en occupe sûrement.

Autre variation sur le même thème, la pratique qui consiste à transférer *délibérément* à quelqu'un d'autre la responsabilité de la maintenance et de la conservation sur le long terme. Un exemple nous en est offert par l'étude de cas sur le système de dépôt en ligne de la Cour suprême de Singapour, où, parce que la Cour suprême n'est pas certaine d'être en mesure de continuer à gérer l'infrastructure à clé

⁴⁶ Voir par exemple l'étude de cas *Horizon Zero/Zero* Horizon Online Magazine, où le producteur, en réponse à la question *Quelles politiques et/ou méthodes de conservation sont mises en place et comment ?* (question de recherche 19a) indique que « la maintenance du site web sera sous-traitée pour une période de dix ans, même si l'essentiel des interactions (chat room, message boards) seront désactivées » (Brent Lee (2004). « InterPARES 2 Project – Case Study 03 Final Report : Horizon Zero/Zero Horizon Online Magazine and Media Database », 9, http://www.interpares.org/display_file.cfm?doc=ip2_cs03_final_report.pdf. De même, dans l'étude de cas consacrée à l'artiste Stelarc, en réponse à la même question de recherche, le producteur déclare que « chaque performance génère sa propre documentation ; chaque dessin, photo, vidéo est converti en cas de besoin dans des formats numériques adaptés, et mis ensuite en ligne sur le site » (Daniel et Payne, « Caste Study 02 Final Report », op. cit., 14).

⁴⁷ Un exemple particulièrement révélateur en est offert dans l'étude générale n° 9, où on peut lire que « prévaut « l'état d'esprit de l'intermédiaire », selon lequel le participant [à l'enquête] (en particulier les consultants) considère la conservation sur le long terme comme le problème exclusif de ses clients. » (Preston, « General Study 09 Final Report », op. cit., 54). Comme l'a exprimé sans état d'âme une des personnes interrogées dans l'étude : « Nous fournissons des données à nos clients et ce qu'ils en font est leur problème, » (Ibid.) oubliant apparemment que, dans un environnement numérique, la conservation commence avec le producteur et s'étend jusqu'au service d'archives et que c'est donc une responsabilité continue et *répartie*.

publique qu'elle utilise pour émettre des certificats numériques aux avocats, elle envisage d'externaliser la gestion de tout ce qui concerne l'infrastructure à clé publique à une autorité de certification agréée afin de se « défaire de la gestion de l'obsolescence technologique⁴⁸ ».

Un autre motif couramment invoqué pour ne rien faire est le coût de la conservation. Mais ces prétendues préoccupations dissimulent en fait souvent en partie, particulièrement dans les organisations, un manque de volonté à apporter des réponses au problème de la conservation et une réticence à permettre aux individus qui s'en préoccupent d'accéder aux moyens de conservation. De nombreux scientifiques, en particulier, ont dénoncé le manque flagrant d'entrepôts adaptés à la conservation des données scientifiques. Par exemple, comme cela a été déjà mentionné, de nombreux participants de l'étude générale sur les pratiques d'archivage électronique des archéologues utilisant un SIG ont laissé libre-cours à leur frustration (et parfois, comme l'indique le rapport, à leur désespoir) quant au manque récurrent d'entrepôts appropriés à la conservation sur le long terme des données et documents archéologiques⁴⁹. Un point de vue un peu différent a été exprimé par le producteur dans l'étude de cas sur l'Atlas cyber-cartographique de l'Antarctique, qui a indiqué que « par-delà des obligations scientifiques et professionnelles et la revue de leurs travaux par des pairs⁵⁰, les producteurs n'ont guère le souci de les conserver, » suggérant par là que, dans de nombreux cas, les chercheurs sont tout aussi responsables du manque de volonté actuel de mener une politique de conservation efficace des données et des documents de la recherche scientifique sur le long terme. Pour faire évoluer les choses, il suggère d'inciter les « agences et organismes qui financent la recherche... à inclure la conservation dans leurs structures d'attribution et aussi... à apporter un support institutionnel et... à réclamer des plans d'action⁵¹. » Enfin, la réticence des individus comme des organisations, en particulier dans la sphère scientifique, à admettre leur rôle dans le cycle de la conservation tient en partie à la nature même des documents qui doivent être conservés, tant leur volume que leur complexité interne et externe. De fait, les ensembles de données, et les documents et documents d'archives associés, sont parfois si volumineux, complexes (par exemple, comprenant du multimédia, de l'interactivité, etc.) et éparpillés (chez un même producteur et entre producteurs) que rares sont les organisations disposées à assumer la responsabilité de leur conservation authentique sur le long terme.

Cette situation explique les efforts déployés par le producteur dans l'étude de cas sur l'Atlas de

⁴⁸ Goh, « Case Study 21 Final Report », op. cit., 33.

⁴⁹ Preston, « General Study 09 Final Report », op. cit., 90. Voir également en page 54 les commentaires de certains participants concernant ce problème.

⁵⁰ Lauriault et Hackett, « Case Study 06 Final Report », op. cit., 31.

⁵¹ Ibid.

l'Antarctique, qui travaille avec la bibliothèque de Carleton University « pour essayer d'archiver le CAA [Atlas de l'Antarctique], tel qu'il existe à la fin du projet, conformément aux exigences du SSHRC [Social Sciences and Humanities Research Council of Canada], » car « pour l'instant, aucune institution n'est en mesure d'archiver les données et les résultats des projets soutenus par le SSHRC⁵². » Qui plus est, bien que l'on espère que le Scientific Committee on Antarctic Research (SCAR) prendra à plus ou moins longue échéance la responsabilité de la maintenance sur le long terme du CAA et de son développement, le producteur reconnaît que ce n'est absolument pas certain, au vu des moyens humains, financiers, institutionnels et techniques limités du SCAR⁵³.

Toujours dans le même ordre d'idées, il arrive que la numérisation soit considérée comme une solution de conservation. Ce point de vue est généralement le fait de producteurs qui : (a) ont un problème matériel d'espace de stockage, (b) ont du mal à trouver certains éléments parmi leurs documents analogiques, (c) se sentent obligés (et y sont parfois tenus par la loi) d'accéder à la demande du public pour un accès électronique distant aux documents d'archives et autres ressources numériques⁵⁴ ou (d) considèrent la numérisation comme un moyen de permettre la conservation grâce un accès étendu aux ressources numérisées et à leur diffusion⁵⁵. Pour beaucoup, la technologie apparaît comme la solution à ces différents problèmes. En règle générale, ces affirmations émanent de producteurs peu au fait des difficultés techniques et des coûts récurrents associés à la conservation d'objets numériques sur le long terme.

Pour ce qui est de l'authenticité, la plupart des producteurs de documents ne voient guère de nécessité à prendre des mesures pour la protéger dans l'environnement numérique. Différentes raisons sont avancées dans les études de cas portant sur des individus et de petits groupes, parmi lesquelles : le manque de valeur monétaire de l'objet⁵⁶, le nombre réduit de praticiens dans un domaine où tout le

⁵² Ibid., 18, 29.

⁵³ Ibid., 18.

⁵⁴ (a), (b) et (c) ont été cités comme des arguments en faveur de la numérisation des livres fonciers étudiée dans l'étude de cas sur la numérisation du livre foncier d'Alsace-Moselle (voir Blanchette et al., « Case Study 18 Final Report », op. cit.).

⁵⁵ Ce point de vue est exprimé sans détour par le répondant dans l'étude de cas Arbo Cyber, théâtre (?) qui, tout en admettant que le « groupe théâtral est plus intéressé par la numérisation comme moyen de diffusion », « considère la numérisation comme un meilleur moyen de conservation que le système d'archivage traditionnel » (Cardin, « Case Study 01 Final Report », op. cit., 48).

⁵⁶ Par exemple, on peut lire dans l'étude de cas n° 9 (3), « les documents analogiques sont conservés essentiellement du fait de leur valeur marchande. Les animations ou les autres œuvres d'art individuelles peuvent être vendues dans des galeries d'art ou lors de ventes publiques mais les fichiers numériques n'ont, en l'espèce, pas de valeur ou seulement une valeur éphémère. L'ironie de la chose est que les fichiers numériques peuvent prendre de la valeur s'ils sont « touchés » par une personne connue, par exemple, une célébrité peut écrire sur l'étiquette d'un support amovible même si elle n'a jamais utilisé l'ordinateur avec lequel le fichier a été créé » (Turner et al., « Case Study 09 (3) Final Report », op. cit., 5).

monde se connaît et reconnaîtrait l'œuvre⁵⁷, ou encore la volonté de rendre le contenu librement accessible à d'autres pour être réutilisé, parfois dans un contexte artistique⁵⁸, parfois dans un contexte scientifique⁵⁹. Pour la plupart, les individus et les petits groupes de la sphère artistique semblent se satisfaire de l'idée qu'ils sont les arbitres de l'authenticité pendant leur vie. Par exemple, dans l'installation multimédia interactive de l'étude de cas *The Danube Exodus*, la supervision directe des créateurs sur leurs objets numériques était considérée par tous les sujets de l'étude comme le seul moyen fiable de maintenir et de garantir l'authenticité de ces objets, de sorte que l'authenticité des objets ne pourrait plus être garantie après le décès de leur créateur⁶⁰.

En revanche, dans les organisations, le point de vue le plus répandu parmi les personnes interrogées était que quelqu'un était sans doute chargé de la mise en place de mesures garantissant l'authenticité et la conservation des documents – ce qui était vrai dans certaines études de cas et pas dans d'autres. Comme cela a été indiqué plus haut, certaines organisations continuent à utiliser des systèmes d'archivage papier, en particulier dans les environnements où des questions juridiques liées la valeur probante des documents – en particulier en relation avec les droits de propriété intellectuelle – risquent

D'un autre côté, comme l'indique un peu plus loin le rapport, « d'autres utilisations marketing et promotionnelles reposent sur la réutilisation d'objets numériques pour des entretiens avec des animateurs, des événements, des prix, et ainsi de suite, comme les bonus sur les versions Dvd des films du studio » (Ibid., 8).

⁵⁷ Comme l'œuvre de performance de l'artiste Stelarc qui, par exemple, considère que l'authenticité des objets numériques émanant de ses performances « tient essentiellement à sa position unique au centre de tout le processus et à la nature unique de chaque performance » (Daniel et Payne, « Case Study 02 Final Report », op. cit., 12). Cette conception de l'authenticité est très proche de celle d'authenticité « personnelle » identifiée par l'équipe thématique 2, où « l'authenticité dénote le degré auquel une œuvre d'art manifeste l'individualité et l'essence de son créateur... [de sorte que] l'artiste est l'œuvre d'art, sans médiation des documents. » Ainsi, « la prévalence de cette conception de l'authenticité explique pourquoi de nombreux artistes ne se préoccupent pas d'indiquer explicitement l'identité de leurs œuvres ; pour eux, il est inconcevable que quiconque puisse produire ou produise un art identique au leur. Tout ce que fait un artiste (ou tout ce dont il dirige la création) est authentique, selon cette définition » (Domain 2 Task Force Report, 146).

⁵⁸ À cet égard, l'exemple le plus emblématique est l'étude de cas n° 1, où on peut lire que « le concept de réutilisation est tellement important pour le groupe [Arbo Cyber, théâtre (?)] que les répondants pensent que même après la dissolution du groupe, les artistes doivent pouvoir utiliser les documents dans des projets individuel ou collectifs futurs. En ce sens, les répondants considèrent la *Ludosynthèse* non seulement comme un témoignage des performances passées mais aussi comme une source d'information pour de futures utilisations. Elle crée de la mémoire » (Cardin, « Case Study 01 Final Report », op. cit., 23). De fait, une des finalités premières de la *Ludosynthèse* – le site web interactif et dynamique du groupe – est de permettre aux « spectateurs-utilisateurs de créer ou recréer dans le style d'Arbo en utilisant des médias numériques » (Ibid., 3).

⁵⁹ Dans un certain sens, la base de données sur le Traité de l'Antarctique, dont « la finalité est de faciliter la découverte de connaissances sur les politiques et les stratégies qui promeuvent « la coopération internationale » et « l'utilisation de l'Antarctique à des fins exclusivement pacifiques » [...] « dans l'intérêt du genre humain », comme l'indique le *Préambule du Traité de l'Antarctique de 1959*, (Berkman et al., « Case Study Final Report », op. cit., 9, 34), relève également de cette catégorie, dans la mesure où la base de données sert de portail à des documents publics dont l'authenticité, insiste le créateur de la base de données, « ne peut être garantie que par les agences gouvernementales qui émettent les documents » (Ibid., 29).

⁶⁰ Sally Hubbard (2006), « InterPARES 2 Project – Case Study 10 Final Report : *The Danube Exodus* », 8. Voir : http://www.interpares.org/display_file.cfm?doc=ip2_cs10_final_report.pdf.

de se faire jour à une date ultérieure⁶¹.

InterPARES 1 avait souligné l'importance de tenir à jour les métadonnées pour tout ce qui concerne la protection de l'identité et de l'intégrité des documents d'archives, par exemple les questions de propriété intellectuelle et de copyright. Cette conclusion a été réaffirmée avec force par InterPARES 2⁶². Certaines études de cas de la sphère artistique témoignent d'une variante intéressante du rôle des métadonnées. Comme pour l'authenticité, un producteur individuel peut seulement contrôler la représentation de ses œuvres pendant sa vie. Pour que l'intentions du créateur puisse lui survivre, et être utilisée pour représenter ses œuvres après sa mort, les paramètres de cette intention doivent être enregistrés et transmis, vraisemblablement dans des métadonnées⁶³. Lorsqu'on a leur demandé d'identifier, parmi une liste de six options, la méthode qui leur semblait la meilleure pour représenter l'identité d'œuvres musicales dépourvues de partition (au sens traditionnel), près d'un quart (25 %) des personnes participant à l'enquête sur les pratiques d'archivage des compositeurs de l'étude générale n° 4 ont choisi l'enregistrement audio ou vidéo de la performance⁶⁴, suggérant par là qu'enregistrer une performance conforme à la vision du compositeur pourrait également être un moyen de fixer son intention.

La politique de conservation la plus prometteuse et la plus avant-gardiste rencontrée par les chercheurs du thème 3 est sans doute celle mise en œuvre dans l'étude de cas scientifique n° 8 (NASA). Ici, le producteur affirme avoir mis en place des procédures de conservation permettant la *conservation par objet dans un format pérenne* (POP) qui est « une technique pour garantir que les documents d'archives électroniques restent accessibles en les convertissant dans des formats auto-descriptifs, indépendants de matériels ou de logiciels spécifiques⁶⁵. » Dans le cas de la NASA, cette démarche repose en partie sur

⁶¹ Voir par exemple l'étude de cas *HorizonZero* – où sont analysées les activités d'archivage d'un magazine d'art électronique – dans laquelle il est noté que les courriels, les contrats et les autres documents juridiques sont conservés sous forme analogique dans un système de classement papier en dehors du système d'« archivage » numérique du producteur où tous les autres documents et objets numériques produits par *HorizonZero* sont conservés (Lee, « Case Study 03 Final Report », op. cit., 4, 6).

⁶² Voir l'étude de cas n° 6 (Atlas cyber-cartographique de l'Antarctique) pour un excellent exemple (Lauriault et Hackett, « Case Study 06 Final Report », op. cit.).

⁶³ La nécessité d'un tel processus a été clairement démontrée par la tentative mentionnée plus haut de ressusciter et représenter une performance authentique d'une œuvre acoustique, dont l'échec partiel a conduit les chercheurs de l'équipe thématique 2 à conclure que « les producteurs, tant qu'ils sont vivants, sont les meilleurs arbitres de l'authenticité des performances. Il leur appartient donc de fournir des descriptions de leurs œuvres qui soient indépendantes des technologies et dont l'authenticité puisse être préservée, afin d'en permettre des performances authentiques à l'avenir (Domain 2 Task Force Report, 156).

⁶⁴ Longton, « General Study 04 Final Report », op. cit., 4 (voir question 11).

⁶⁵ Richard Pearce-Moses, « Persistent Object Preservation », in *A Glossary of Archival and Records Terminology* (Chicago : The Society of American Archivists, 2005). Voir : <http://www.archivists.org/glossary/index.asp>. L'utilisation de « formats auto-descriptifs » (c'est-à-dire, POP ou étiquetage) est une des douze politiques élémentaires de

l'utilisation de l'Object Description Language⁶⁶ (ODL, Langage de description orienté objet) pour « créer des étiquettes (descriptions de données) pour des fichiers de données et d'autres objets comme les logiciels et les documents⁶⁷. » En cas d'obsolescence technologique, il suffit avec cette méthode de migrer l'interpréteur pour le langage de description de la structure des fichiers ; il n'est pas nécessaire de créer un logiciel d'accès pour chaque structure de fichier⁶⁸. Depuis le déploiement du Planetary Data System (PDS), il y a plus de quinze ans, « aucun des produits de données n'a nécessité de mise à jour (conversion ou migration) dans d'autres formats de données⁶⁹ », ce qui suggère que cette politique pourrait offrir une solution viable pour un des aspects au moins de la conservation sur le long terme des documents numériques⁷⁰.

Réponses au questionnaire de recherche

Question 1

Les concepts, méthodes et modèles d'évaluation élaborés par InterPARES 1 pour les documents d'archives administratifs et juridiques produits dans des bases de données et des systèmes de gestion des documents s'appliquent-ils à l'évaluation des documents d'archives produits par des activités artistiques, scientifiques et administratives résultant de l'utilisation des technologies étudiées par InterPARES 2⁷¹ ?

conservation de base (B1.1) décrites dans une annexe aux *Principes directeurs à l'usage des producteurs* élaborés par les chercheurs de l'équipe thématique 3 (voir l'annexe 21). Voir également l'annexe 21c, « Politiques de maintenance et de conservation des documents d'archives numériques ».

⁶⁶ L'ODL est un langage créé par la NASA qui est utilisé pour encoder les étiquettes de données de son Planetary Data System (PDS) et d'autres systèmes de données (voir National Aeronautics and Space Administration, « Chapter 12. Object Description Language Specification and Usage » in *Planetary Data System Standards Reference*, version 3.7., 20 mars 2006. JPL D-7669, Part 2 (Pasadena, CA : Jet Propulsion Laboratory, California Institute of Technology, 2006). Voir : <http://pds.nasa.gov/documents/sr/>.

⁶⁷ William Underwood (2005), « InterPARES 2 Project – Case Study 08 Final Report : Mars Global Surveyor Data Records in the Planetary Data System », 15. Voir : http://www.interpares.org/display_file.cfm?doc=ip2_cs_08_final_report.pdf.

⁶⁸ Ibid., 29.

⁶⁹ Ibid.

⁷⁰ Il est intéressant de souligner que, sur la base des succès réitérés des tests de la politique POP par le Distributed Object Computation Testbed (DOCT), collaboration entre l'Advanced Research Projects Agency du Département de la défense américain, l'U.S. Patent and Trademark Office et l'U.S. National Archives and Records Administration (NARA), la NARA considère l'approche POP « comme la plus prometteuse jamais proposée pour conserver des informations numériques en général et des documents d'archives électroniques en particulier » (Kenneth Thibodeau (2001). « Building the Archives of the Future », *D-Lib Magazine* 7(2). Voir :

<http://www.dlib.org/dlib/february01/thibodeau/02thibodeau.html>) et a choisi cette méthode de conservation pour

l'Electronic Records Archives (ERA) qui doit voir le jour prochainement (voir : <http://www.archives.gov/era/about>).

⁷¹ Projet InterPARES 2, page web du thème 3. Voir : http://www.interpares.org/ip2/ip2_domain3.cfm.

Les études de cas d'InterPARES 2 montrent que les difficultés auxquelles se heurtent traditionnellement les services d'archives pour évaluer les documents se posent avec plus d'acuité dans l'environnement étudié. Absence de métadonnées, modes de classement particuliers et éparpillement des documents n'en sont que quelques exemples. Plus étonnant, sans doute, est le constat que les documents numériques, et en particulier les documents non textuels, demeurent dans certains cas très dépendants du matériel informatique, une situation de plus en plus rare dans les environnements d'archivage des documents textuels des organismes publics et des entreprises étudiés par InterPARES 1. Certaines études de cas ont également révélé l'absence de critères de sélection des documents [en vue de leur conservation] mais la définition de ces critères ne fait pas partie des fonctions du service d'archives touchées par la technologie numérique, si l'on excepte l'importance nouvelle de la fonctionnalité pour certains types de documents numériques, comme ceux contenus dans des applications de bases de données.

Les études de cas d'InterPARES 1 avaient également clairement démontré la nécessité d'indiquer, au cours de l'évaluation, si les systèmes considérés contenaient ou pouvaient produire des documents d'archives. Le fait demeure pour les applications interactives, expérientielles et dynamiques étudiées par InterPARES 2, bien que le processus puisse s'avérer plus complexe étant donné le nombre élevé de composants et la complexité croissante de leurs relations.

L'exploration d'exemples interactifs, dynamiques et expérientiels dans le cadre d'InterPARES 2 a considérablement enrichi la réflexion sur la nature des documents d'archives numériques issue d'InterPARES 1. Les implications des résultats des études de cas d'InterPARES 2 pour la définition du document d'archives font l'objet d'une étude approfondie dans l'article « Le concept de document d'archives dans les environnements interactifs, expérientiels et dynamiques : le point de vue d'InterPARES »⁷². Une des conclusions majeures d'InterPARES 1 était que « il n'est pas possible de préserver un document d'archives numérique ; seule peut être préservée la capacité de le reproduire⁷³. » Cette conclusion reposait sur l'idée qu'il n'est pas possible qu'un document d'archives stocké sous forme binaire (et donc non lisible pas l'être humain) puisse servir de document d'archives. Les conclusions issues d'InterPARES 2, fondées cette fois sur l'étude de systèmes électroniques interactifs, expérientiels et dynamiques, jointes à la réaffirmation de la définition centrale du « document d'archives » et étayées par des précédents historiques, démontrent que cette conclusion n'est valide que

⁷² Duranti et Thibodeau, « Le concept de document d'archives », op. cit.

⁷³ Thibodeau et al., « Preservation Task Force Report », op. cit., 106.

pour un sous-ensemble de documents numériques, à savoir, ceux qui, pour être effectifs, doivent être représentés à un être humain.

Les études de cas d'InterPARES 2 dans le domaine artistique et le domaine scientifique (en particulier la fabrication) ont identifié des *documents d'archives instrumentaux* conservés dans, ou associés à, des applications ou des systèmes informatiques, qui ne sont pas destinés à produire un document lisible par l'être humain. Les instructions et les spécifications informatiques qui rendent possible l'exécution d'œuvres musicales numériques ou qui pilotent la fabrication assistée par ordinateur de pièces en constituent des exemples évidents. Leur finalité n'est pas la (re)production d'un document lisible par l'être humain mais la (re)exécution ou la (re)production assistées par ordinateur. La définition du document d'archives n'impose pas de contraintes a priori sur le contenu ou la forme du document. Pour déterminer si un document numérique⁷⁴ est un document d'archives, il faut examiner si le document contient et présente les éléments et attributs nécessaires du document d'archives, à savoir : participe-t-il à une action, a-t-il un lien archivistique, a-t-il un auteur, un rédacteur et un destinataire, etc⁷⁵. Tout document numérique satisfaisant ces conditions requises est un document d'archives, même s'il peut être totalement impossible à un être humain de comprendre le document dans sa forme numérique.

Les documents d'archives numériques sont composés d'un ou plusieurs objets, que les chercheurs d'InterPARES 1 ont appelé des « composants numériques^{76*} ». Lorsqu'ils sont « stockés » (c'est-à-dire, encodés sous forme numérique et placés dans un système de stockage sur un support numérique), ces composants numériques ont nécessairement une forme externe différente de celle d'un document numérique dans une forme compréhensible par l'être humain (c'est-à-dire, un document représenté). Poussant l'exploration plus loin, les travaux d'InterPARES 2 ont révélé que, s'ils sont conservés et

⁷⁴ La base de données terminologique d'InterPARES 2 définit le « document » comme « Une unité d'information indivisible constituée d'un message fixé sur un support (enregistré) d'une manière syntaxique stable. Un document a une forme fixe et un contenu stable », alors que le « document numérique » est défini comme « Un composant ou groupe de composants numériques, qui est sauvegardé [c'est-à-dire, fixé sur un support numérique de stockage non volatile] et traité et géré comme un document. » Un « objet numérique » est l'« Agrégation discrète d'un ou plusieurs trains de bits et des métadonnées de propriétés de l'objet et, le cas échéant, des méthodes pour réaliser des opérations sur l'objet. »

⁷⁵ Pour un développement détaillé sur les éléments et les attributs du document d'archives, voir le *Template for Analysis* d'InterPARES 1, en Annexe 1 de Duranti (éd.), *The Long-term Preservation of Authentic Electronic Records*, op. cit., 192-203. Une version électronique est disponible à l'adresse suivante : http://www.interpares.org/book/interpares_book_i_app01.pdf.

⁷⁶ Le glossaire d'InterPARES 1 en donne la définition suivante : « Objet numérique faisant partie d'un document électronique, ou d'un document électronique reproduit, ou qui contient un ou plusieurs documents électroniques, et qui requiert des méthodes de conservation spécifiques» (http://www.interpares.org/book/interpares_book_q_gloss.pdf).

* *NdT* : cette définition a été modifiée au cours du projet. Voir la base de données terminologique d'InterPARES 2 : http://www.interpares.org/ip2/ip2_terminology_db.cfm et la traduction française du glossaire. Voir également le texte un peu plus loin.

gérés comme un document d'archives, le ou les composants numériques stockés d'un document d'archives numérique représenté constituent collectivement un document numérique stocké. Par conséquent, InterPARES 2 a distingué deux sous-catégories de documents d'archives numériques : le « document d'archives numérique stocké », c'est-à-dire, l'encodage d'un document numérique dans un système⁷⁷ ; et le « document d'archives numérique manifesté », c'est-à-dire, un document d'archives numérique stocké qui est visualisé dans une forme pouvant être présentée à un être humain (c'est-à-dire, dans une forme lisible par l'être humain) ou à un système informatique (c'est-à-dire, en langage machine). Concrètement, un document d'archives numérique stocké est ce qui est conservé, alors que l'affichage ou toute autre « sortie » d'un document numérique représenté à partir du document numérique stocké, est une copie qui est reproduite à la demande dans le but de communiquer des informations à des personnes ou à des systèmes informatiques, et non comme un document qui est conservé. Ainsi, dans ce type de cas, un document d'archives numérique stocké peut être considéré comme un document d'archives parce qu'il est destiné à être utilisé, et peut être utilisé, pour reproduire un document d'archives numérique représenté⁷⁸.

L'analyse des études de cas d'InterPARES 2 a montré qu'à un document numérique stocké ne correspond pas systématiquement un document numérique représenté. Un document numérique stocké (ou l'un quelconque de ses composants) peut être utilisé avec – ou, comme dans le cas des documents instrumentaux, peut même contrôler – les saisies et les interactions de l'utilisateur et des données variables d'autres sources pour produire des représentations variables qui ne peuvent être des documents d'archives précisément parce que leur contenu et/ou leur forme ne sont pas fixes. En conséquence, un document d'archives numérique stocké est considéré comme un document d'archives

⁷⁷ Bien qu'InterPARES 2 définisse littéralement un document d'archives numérique stocké comme « Un document numérique qui est traité et géré comme un document d'archives », lorsqu'on prend en compte le fait qu'un document numérique est composé d'un objet numérique, la définition devient « Un objet numérique, placé dans un système de stockage sur un support numérique, qui est traité et géré comme un document d'archives, et qui contient des informations sur les propriétés de l'objet et peut également contenir des méthodes pour réaliser des opérations sur ou avec l'objet. » Bien que cela semble concerner plus directement la tâche de conservation que celle d'évaluation, il est essentiel d'admettre qu'un document d'archives numérique stocké ne correspond pas nécessairement à un seul fichier stocké sur un support numérique, mais plutôt à un amalgame de composants numériques discrets, ou de sous-ensembles de composants, logiquement liés les uns aux autres au niveau du système de fichiers par les informations encodées sous forme numérique dans le bloc d'en-tête du fichier, et au niveau de l'application, par des données de contenu, des données de forme, des données de composition et les règles associées, encodées sous forme numérique. Dès lors, les composants numériques d'un document d'archives numérique stocké peuvent être un sous-ensemble d'un fichier stocké, un ensemble de fichiers stockés ou différents éléments de données extraits de différents fichiers.

⁷⁸ Dans la plupart des cas, mais pas dans tous, rendre possible la reproduction d'un document d'archives numérique représenté est le but premier de la conservation du document d'archives numérique stocké, alors que le document d'archives représenté est reproduit pour communiquer des informations à des personnes ou à d'autres systèmes (Duranti et Thibodeau, « Le concept de document d'archives », op. cit., 51).

parce qu'il satisfait à la définition du document d'archives, qu'il participe ou non à la production d'autres documents ou qu'il fournisse ou non un moyen de reproduire un document d'archives numérique représenté.

InterPARES 2 a simplifié la définition du « composant numérique » qu'avait retenue InterPARES 1, la rendant à la fois moins ambiguë et plus facile à utiliser : « Agrégation de données encodées sous forme numérique composée d'un ou plusieurs trains de bits, y compris les métadonnées nécessaires pour ordonner, structurer ou représenter son contenu et sa forme, requérant une action de conservation précise^{79*}. » Les documents et documents d'archives numériques stockés sont constitués de trois types de données : (1) les données de contenu, qui se rapportent aux actions, aux faits ou aux données que le document ou le document d'archives véhicule (c'est-à-dire, son contenu) ; (2) les données de forme, qui permettent au système de reproduire le document ou le document d'archives dans la forme correcte à partir du document ou document d'archives numérique stocké ; et (3) les données de composition, qui identifient les éléments de données de forme et de contenu stockées qui appartiennent au document ou au document d'archives et les associent aux différents éléments de structure définis par les données de forme. Un composant numérique peut être constitué d'un ou plusieurs de ces types de données. Sont également considérées comme un type important de composants numériques les règles encodées sous forme numérique qui gouvernent la reproduction du contenu et de la forme du document ou document d'archives numérique représenté en déterminant ses caractères externes, ainsi que toute variation admissible de forme ou de contenu.

InterPARES 2 a également proposé d'élargir la typologie des documents d'archives (dispositif, probant, à l'appui ou narratif). Le nouveau type, le *document d'archives prospectif*, ne se rapporte pas à une action ou un état passés mais détermine la forme et/ou le contenu de documents d'archives, actions ou états produits ultérieurement. Les documents prospectifs se subdivisent en documents *d'instructions* et documents *instrumentaux*. Un document d'archives d'instructions est un document qui fournit des instructions, destinées à être lues par l'être humain, sur la production de documents dans le cours d'une activité. Il peut exister sous forme papier ou comme document numérique représenté. Inversement, un document instrumental est un document qui est utilisé, dans la forme numérique sous laquelle il est stocké, par un système informatique pour exécuter une action, une interaction ou un processus, et qui ne peut remplir sa fonction s'il est converti dans un format lisible par l'être humain. Pour être

⁷⁹ Base de données terminologique d'InterPARES 2. Voir : http://www.interpares.org/ip2/ip2_terminology_db.cfm.

* *NdT* : la définition donnée dans le texte est différente de celle proposée dans la base de données terminologique d'InterPARES 2.

considérés comme des documents d'archives, les documents instrumentaux doivent être « correctement maintenus et gérés [dans leur forme numérique stockée] comme des éléments intellectuellement liés d'ensembles de documents d'archives⁸⁰. »

Dès lors, il convient de clarifier les conclusions d'InterPARES 1. Comme cela déjà été évoqué, le composant numérique, ou le jeu de composants numériques, utilisé pour reproduire une copie authentique d'un document lisible par l'être humain peut lui-même être un document d'archives ; à savoir, un document d'archives numérique stocké. Cependant, il est important de comprendre qu'un document d'archives numérique stocké ne peut être utilisé pour la finalité pour laquelle le document numérique représenté est reproduit. En effet, une étape au moins le sépare de cet état. De nombreuses étapes peuvent être requises pour traiter les composants numériques en vue de générer le document numérique représenté dans la forme lisible par l'être humain requise. En conséquence, on peut dire que la finalité du document d'archives numérique stocké n'est pas de remplir la finalité à l'appui, narrative ou de dispositif du document d'archives (lisible par l'être humain) mais de rendre possible la production de copies authentiques de ce document. Pour valide qu'elle soit, cette conception des choses semble par trop restrictive. Quand, par exemple, un étudiant s'inscrit en ligne à un cours, en supposant que le système d'inscription de l'université soit un système d'archivage digne de confiance, les données du dossier d'inscription de l'étudiant qui est créé à partir de cette transaction – et qui est conservé dans le système d'archivage comme document numérique stocké – peuvent être utilisées, en combinaison avec d'autres données comparables extraites des documents numériques d'inscription stockés d'autres étudiants, pour produire un nouveau document – une liste de participants – sans passer par l'étape intermédiaire de reproduction des dossiers d'inscription individuels dans une forme lisible par l'être humain (autrement dit, les documents représentés) ou même de reproduction des documents numériques représentés des inscriptions individuelles pour être utilisées par le système informatique d'inscription. Il suffit d'extraire les éléments de données pertinents (le cours et le niveau) des documents numériques stockés correspondants (le document d'inscription de chaque étudiant) pour produire la liste de participants. Lorsqu'elles sont utilisées de cette manière, les versions numériques *stockées* des documents d'inscription, par opposition à leurs versions numériques représentées, doivent être considérées comme les documents d'archives qui ont un lien archivistique avec le document de la liste de participants et, partant, doivent être gérées et évaluées en conséquence.

Les études de cas ont ainsi mis en lumière l'existence de plusieurs nouveaux types de documents,

⁸⁰ Duranti et Thibodeau, « Le concept de document d'archives », op. cit., 59.

certains ayant leur équivalent dans l'univers analogique et d'autres ne pouvant exister que dans un environnement numérique. Ces documents numériques sont produits dans des systèmes de documents avec des groupes de plus en plus complexes de composants numériques, multipliant le nombre de relations qui doivent être identifiées, décrites, gérées et conservées. Certains systèmes ne peuvent afficher leurs documents numériques stockés dans une forme lisible par l'être humain. Il y a aussi des systèmes qui peuvent afficher le contenu d'un document numérique représenté dans plusieurs formes documentaires (comme les données de tableur), bien que chaque forme disponible soit en fait fixe et stable. Il y a des systèmes où les documents non lisibles par l'être humain participeront à la production de contenu numérique qui ne sera jamais fixé et, par conséquent, n'est pas un document d'archives. Et il y a des systèmes qui semblent produire du contenu variable, alors qu'une analyse plus minutieuse confirme que ce contenu apparemment mouvant est borné par des règles fixes et est donc, de fait, fixe et stable. Appelé « variabilité limitée » dans le domaine des sciences de l'information, ce concept désigne un environnement où « les changements de la forme et/ou du contenu d'un document d'archives numérique... sont limités et contrôlés par des règles fixes [et un ensemble stable de données de contenu, de données de forme et de données de composition], de sorte que la même recherche, requête ou interaction produit toujours le même résultat⁸¹. » Il est important de souligner qu'il s'agit là d'un aspect des documents numériques, avec des conséquences pour l'évaluation, qui va au-delà de la conclusion d'InterPARES 1 (réaffirmée par InterPARES 2) selon laquelle, bien que l'intégrité matérielle d'un document – nombre de chaînes de bits par exemple – puisse être compromise, aussi longtemps que le document conserve tous ses attributs essentiels de telle sorte que le message qu'il est censé communiquer pour accomplir sa finalité n'est pas modifié, il peut être considéré « complet et inchangé⁸². »

En fait, à travers le concept de variabilité limitée, InterPARES 2 a élargi les circonstances dans lesquelles des variations de la forme et du contenu d'un document peuvent être considérées comme acceptables, par-delà les considérations élémentaires d'altérations involontaires, accidentelles ou fortuites de l'intégrité matérielle du document, pour prendre en compte les intentions du producteur concernant l'utilisation d'éléments de présentation variables des documents. Relativement aux variations de la forme documentaire, il est important d'admettre que :

⁸¹ Tiré de la définition de la « variabilité limitée » de la Base de données terminologique d'InterPARES 2. Voir : http://www.interpares.org/ip2/ip2_terminology_db.cfm. Pour une présentation plus détaillée du concept, voir Duranti et Thibodeau, « The Concept of Record », op. cit., 47-49.

⁸² MacNeil et al., « Appraisal Task Force Report », op. cit., 47.

Dans de nombreux documents interactifs, expérientiels et dynamiques, les auteurs ou les rédacteurs ont délibérément recours aux possibilités spécifiques offertes par la technologie numérique quant à la variabilité de la forme dans laquelle les informations sont présentées. Dans ce type de cas, la forme est « fixe » en ce sens que la conception autorise la variation de certains aspects de forme et pas d'autres. Les formes documentaires qui incluent des éléments variables ne contreviennent pas aux exigences de forme fixe, pas plus que des enregistrements audio ou vidéo analogiques, qui présentent des variations du son et de l'image dans le temps. Cette variabilité de la présentation voulue par l'auteur doit être considérée comme faisant partie des caractères externes de la forme documentaire... Avec les documents numériques, dès lors, la forme « fixe » désigne les aspects de forme que l'auteur ou le rédacteur veut ou peut contrôler⁸³.

De la même manière, un producteur peut provoquer ce type de « variabilité limitée » intentionnelle dans le contenu d'un document sans affecter l'intégrité de celui-ci. Par exemple, comme c'est le cas de certains catalogues de vente en ligne, les environnements numériques interactifs donnent la possibilité à un producteur d'utiliser des formes documentaires qui autorisent une sélection variable du contenu et un ordre variable de cette sélection, comme dans l'affichage de sous-ensembles du contenu de ces catalogues en réponse à une action donnée de l'utilisateur⁸⁴. Pourvu que le *document d'archives numérique stocké* soit contrôlé par des règles fixes telles que la même recherche, la même requête ou la même interaction génère toujours la même forme documentaire et la même sélection de contenu dans le *document d'archives numérique représenté*⁸⁵, ces cas peuvent satisfaire la condition requise de contenu fixe⁸⁶.

Une autre différence qui doit être prise en compte lors de l'évaluation des documents dans les environnements numériques interactifs, expérientiels et dynamiques est que l'utilisation de la technologie peut modifier les activités du producteur. Changements qui se traduiront à leur tour par des

⁸³ Duranti et Thibodeau, « Le concept de document d'archives », op. cit., 47-48.

⁸⁴ Ibid., 49.

⁸⁵ Défini dans la Base de données terminologique d'InterPARES 2 comme « Un document numérique qui est visualisé ou restitué à part d'un document numérique stocké et/ou de composants numériques stockés dans une forme adaptée à la présentation à un être humain (c'est-à-dire, dans un format lisible par l'être humain) ou à un système informatique (c'est-à-dire, en langage machine). »

⁸⁶ Il est également important de souligner que dans ce type de cas, alors que le document numérique représenté semble être un document d'archives, c'est en fait le document d'archives numérique stocké – qui englobe toute la « palette de possibilités » du catalogue de vente (Duranti et Thibodeau, « Le concept de document d'archives », op. cit., 32) vis-à-vis de la forme et du contenu de son document d'archives, pas seulement le sous-ensemble présenté à l'utilisateur en réponse à une requête – qui est le document d'archives qui est conservé et utilisé pour consultation future. Duranti et Thibodeau qualifient ces types de documents numériques représentés qui semblent être des documents d'archives mais ne sont en réalité pas représentés au niveau interne par un document d'archives stocké qui correspond exactement au document d'archives numérique représenté apparent, de pseudo-documents (Ibid., 27).

changements au niveau des activités de production et d'archivage des documents. La nécessité, en particulier, de réexaminer périodiquement les activités de production (et les technologies associées) des documents et des documents d'archives du producteur après l'évaluation initiale, transparaît ainsi clairement dans l'étude de cas VanMap. Bien que le système d'information géographique existant ne soit pas conçu pour produire des documents d'archives, il est envisagé de modifier son architecture pour que cela devienne possible. À la différence des conclusions d'InterPARES 1, qui avaient mis l'accent sur le suivi des seuls documents d'archives sélectionnés en vue de leur conservation définitive, cette observation plaide également en faveur de la nécessité, soulignée par les chercheurs d'InterPARES 2, d'étendre la fonction de surveillance pour y inclure les données et les documents que des évaluations antérieures avaient exclus de la conservation, ainsi que les systèmes et les activités dont des évaluations antérieures avaient estimé qu'ils ne produisaient pas de documents d'archives. Cependant, même lorsque les documents eux-mêmes (ou les systèmes qui les produisent) ne changent pas, leur mise en ligne peut être à l'origine de changements importants du lien archivistique, la facilité d'accès se traduisant souvent par une utilisation des documents dans davantage d'activités qu'auparavant. Dans la mesure où l'évaluation doit sélectionner les documents à conserver en se fondant sur la connaissance des processus du producteur et de la totalité de son fonds d'archives, l'évaluation dans les environnements numériques ne doit pas être considérée comme une opération unique sur un ensemble statique de documents mais comme une activité qui doit s'étendre sur toute la durée de vie des documents et englober les activités de leurs producteurs depuis la production des documents jusqu'à leur versement au services d'archives en charge de leur conservation définitive⁸⁷. Il va sans dire que cela rend le processus d'évaluation beaucoup plus complexe, en particulier par rapport aux évaluations « en une seule fois » généralement appliquées aux documents analogiques traditionnels.

Enfin, une dernière différence entre le processus d'évaluation envisagé par InterPARES 1 et les conclusions d'InterPARES 2 concerne la différenciation plus marquée qui existe désormais, dans un environnement numérique, entre les documents rétrospectifs et les documents prospectifs, en particulier ceux qui ont un rôle instrumental dans la (re)production d'un autre document. Outre que préserver cette

⁸⁷ Ce processus est englobé dans l'activité A4.2.4 (Suivi des décisions d'évaluation) du modèle de Chaîne archivistique, que le modèle définit ainsi : « Contrôler les décisions d'évaluation relativement aux évolutions ultérieures intervenant dans les activités du producteur et/ou du service d'archives qui pourraient nécessiter d'ajuster ou de refaire l'évaluation, comme des changements importants dans : (1) les documents évalués et/ou leur contexte, (2) la mission et les responsabilités du producteur, (3) les activités ou systèmes de gestion et d'archivage des documents du producteur, (4) les activités ou systèmes de conservation des documents du service d'archives et/ou (5) la mission et les responsabilités du service d'archives » (Appendix 14, 224).

fonction de reproduction soulève des difficultés technologiques supplémentaires, le service d'archives doit déterminer si ce type de documents relève de la mission de son institution. Les documents d'archives instrumentaux ne sont pas un concept totalement inconnu des pratiques de conservation d'hier : les partitions permettent des interprétations futures ; les négatifs photographiques et les pellicules de films produisent de nouvelles copies ; les composants numériques interagissent pour produire des copies authentiques des documents. Pour ce qui est des documents instrumentaux, cependant, la conservation ferait d'avantage figure d'obligation « prospective » vis-à-vis du document que d'obligation rétrospective.

Questions 2 et 3⁸⁸

Les concepts, méthodes et modèles de conservation élaborés par InterPARES 1 pour les documents administratifs et juridiques produits dans des bases de données et des systèmes de gestion des documents s'appliquent-ils à la conservation des documents d'activités artistiques, scientifiques et administratives résultant de l'utilisation des technologies examinées par InterPARES 2 ?

Quels paradigmes de conservation peuvent être appliqués à toutes les activités et toutes les technologies ? Quels paradigmes de conservation sont requis pour les types particuliers de documents résultant de chaque activité ?

Comme pour l'évaluation, les études de cas ont mis en lumière nombre des problèmes familiers qui pèsent sur la capacité des services d'archives à protéger les documents sur le long terme. Mauvaises pratiques de production des documents, mauvaise organisation des documents, mauvaises pratiques de maintenance, absence de planification sur le long terme, instabilité du support des documents, mauvaises conditions de stockage et documentation insuffisante ne sont quelques exemples bien connus de l'ère analogique qui demeurent d'actualité. Étant donné la nature de plus en plus complexe, et bien souvent unique des documents d'archives numériques associés aux applications interactives, expérientielles et dynamiques étudiées par InterPARES 2, il va sans dire que nombre des problèmes de conservation hérités de l'univers analogique sont, dans de nombreux cas, beaucoup plus aigus dans l'univers numérique.

Les travaux d'InterPARES 2, particulièrement en ce qui concerne la conceptualisation des documents d'archives dans les systèmes électroniques interactifs, expérientiels et dynamiques, ont entraîné plusieurs changements de paradigmes ayant des incidences pour la conservation numérique, qui

⁸⁸ Projet InterPARES 2, page web du thème 3. Voir : http://www.interpares.org/ip2/ip2_domain3.cfm.

prolongent et élargissent, plutôt qu'ils ne contredisent, ceux d'InterPARES 1. D'abord, alors qu'une « des conclusions les plus importantes d'InterPARES 1 a été de reconnaître et de préciser la différence entre la forme dans laquelle un document électronique est représenté à une personne et la forme dans laquelle il est stocké numériquement », InterPARES 2 a enrichi cette distinction en suggérant que les composants numériques individuels, ou les ensembles de composants numériques, pouvaient eux-mêmes constituer un document d'archives ou un groupe de documents d'archives, selon la façon dont ils sont représentés dans le système et la façon dont ils sont utilisés par le producteur.

Ce constat, nous l'avons dit, a conduit les chercheurs d'InterPARES 2 à distinguer deux sous-catégories fondamentales de document d'archives numériques : les documents numériques stockés et les documents numériques représentés. Bien que le document numérique stocké, constitué d'un ou plusieurs composants numériques, et le document numérique représenté reproduit à partir de ces composants soient apparentés, ce sont des documents distincts et, comme indiqué plus haut dans l'exemple de l'inscription scolaire, ils accomplissent leurs finalités respectives de façons fondamentalement différentes : le premier est utilisé dans sa forme binaire pour « présentation » à d'autres applications ou systèmes électroniques, alors que le second est converti de sa forme binaire en une forme adaptée à la présentation à une personne. Ce fait a des conséquences importantes pour la conservation en ce que les exigences de conservation spécifiques pour les deux types de documents peuvent être différentes – en fonction, notamment, de la façon dont les documents sont utilisés pour accomplir leurs finalités respectives. De fait, un document, comme par exemple un document prospectif instrumental – ou, peut-être aussi, selon les circonstances, un document rétrospectif plus « traditionnel » -, qui a seulement besoin d'être « présenté » et utilisé dans sa forme binaire pour accomplir sa finalité appellera vraisemblablement des mesures de conservation différentes qu'un document rétrospectif ou un document prospectif d'instructions qui, pour accomplir leur finalité, doivent être reproduits dans leur forme numérique représentée pour présentation à un être humain.

Le deuxième changement paradigmatique ayant une incidence pour la conservation concerne, comme cela a été vu plus haut, la différenciation plus marquée qui existe désormais, dans un environnement numérique, entre les documents rétrospectifs et les documents prospectifs, et particulièrement ceux qui jouent un rôle instrumental dans la (re)production d'un autre document. Les documents prospectifs instrumentaux, en particulier, posent des défis de conservation uniques du fait qu'ils doivent impérativement être maintenus dans les systèmes où ils ont été produits – ou dans des systèmes aux fonctionnalités identiques – pour pouvoir préserver leur aptitude à produire ou rendre possibles les

interactions, les expériences, les performances ou tout autre processus qu'ils étaient destinés à produire. Bien que l'encodage du train de bits des composants numériques stockés des documents rétrospectifs et des documents prospectifs d'instructions qui sont conservés pour reproduire des documents numériques représentés puisse, dans de nombreux cas, être converti d'un format dans un autre (par exemple, de .doc en .pdf) sans compromettre l'aptitude du document numérique représenté d'accomplir sa finalité⁸⁹, le train de bits d'un document instrumental doit être conservé dans sa forme originale pour que le document accomplisse sa finalité.

Le troisième changement de paradigme renvoie au concept de variabilité limitée et à son incidence sur notre compréhension des documents numériques représentés. En particulier, ce nouveau concept enrichit le concept de document représenté pour englober tout type de variabilité de forme et de contenu spécifique au document (c'est-à-dire, représentant l'intention de l'auteur) et gouverné par des règles ou des instructions *fixes*⁹⁰.

Ces différents constats ont conduit à s'interroger sur la possibilité de conserver cette gamme plus étendue et plus complexe de documents à l'aide de politiques de conservation établies dans le cadre des paradigmes de conservation existants. Comme cela a été évoqué dans la section Conclusions, les études de cas du projet, centrées sur la production et la maintenance des documents, offraient peu d'exemples concrets de scénarios de conservation, hormis la numérisation et la « mise en ligne ». La discussion précédente sur les variations de forme, de fonction ou même de comportement des documents numériques suggère que la conservation est d'ores et déjà possible à l'intérieur des paramètres connus des politiques de conservation actuelles. Ces documents comportent indéniablement davantage de composants, dont les relations sont d'une complexité croissante. La reproduction exacte et authentique des documents produits par des systèmes interactifs, expérimentiels et dynamiques exigera des métadonnées (et/ou une gestion des métadonnées) de plus en plus élaborées pour documenter l'intention du producteur ou les limitations imposées à cette intention par l'inadéquation des technologies disponibles⁹¹. De nouveaux types de dépendances vis-à-vis du matériel devront être surmontés. De nouvelles méthodes pour confirmer l'identification et la conservation effectives des documents numériques devront être définies, en particulier lorsque la reproduction du contenu et de la

⁸⁹ Voir Duranti et Thibodeau, « Le concept de document d'archives », op. cit., 20.

⁹⁰ Ibid., 51.

⁹¹ Pour une analyse complète de l'existant et des perspectives concernant les métadonnées se rapportant à l'établissement de la fiabilité et de l'authenticité sur le long terme des documents numériques, ainsi qu'à leur conservation et leur ré-exploitation potentielle, voir Description Cross-domain Task Force Report : http://www.interpares.org/display_file.cfm?doc=ip2_book_part_6_description_task_force.pdf.

forme documentaire dans un format lisible par l'être humain (c'est-à-dire, le document numérique représenté) n'est plus le but du processus de conservation.

Du fait de ces difficultés, les systèmes envisagés pour conserver les documents d'archives numériques devront être capables, notamment, de :

1. Identifier et localiser tous les composants numériques de chaque document présent dans le système.
2. Gérer chaque composant numérique conformément à ses exigences de conservation spécifiques.
3. Identifier et gérer comme des documents les composants numériques qui sont utilisés comme des documents, y compris les composants numériques qui sont eux-mêmes des documents (par exemple, les documents instrumentaux) et/ou sont des composant d'un ou plusieurs documents.
4. Identifier et gérer le matériel et les logiciels requis pour traiter tous les composants numériques dans les formats requis pour qu'ils remplissent leur(s) fonction(s) en tant que documents et/ou composants de documents numériques stockés et/ou représentés.
5. Utiliser les logiciels et, le cas échéant, le matériel appropriés à chaque composant pour reproduire le document numérique représenté à partir du document numérique stocké et/ou traiter le document numérique stocké lorsque celui-ci est aussi un document instrumental.

Question 4⁹²

Quelles métadonnées sont nécessaires pour soutenir l'évaluation et la conservation de documents d'archives numériques authentiques résultant de chaque activité ?

On avait espéré dans un premier temps que la réponse à cette dernière question de recherche du thème 3 découlerait naturellement de l'examen des études de cas par les archivistes. Il a fallu se rendre à l'évidence que cela ne serait pas possible, et ce, pour deux raisons. Premièrement, comme cela a été dit, les études de cas étaient centrées sur les producteurs ; or, aucun d'entre eux ou presque n'était impliqué dans l'évaluation de ses documents, ou ne s'en préoccupait, et beaucoup d'entre eux n'étaient pas non plus impliqués dans la conservation de leurs documents sur le long terme. En second lieu, il n'a pas été possible, à cause du nombre limité d'archivistes impliqués dans le projet, d'évaluer des documents des producteurs étudiés dans les études de cas ni, par voie de conséquence, de définir une politique de conservation pour ces documents. Très peu de données ont donc été collectées qui auraient pu être utilisées par les chercheurs du thème 3 pour répondre à cette question. Toutefois, les rapports de

⁹² Projet InterPARES 2, page web du thème 3 : http://www.interpares.org/ip2/ip2_domain3.cfm.

l'équipe transversale Description et de l'équipe transversale Modélisation fournissent des analyses et des discussions détaillées des métadonnées requises pour soutenir l'évaluation et la conservation de documents numériques authentiques. En fait, cette dernière question du thème 3 a été transférée à l'équipe transversale Modélisation et intégrée à ses travaux au cours de l'élaboration du modèle de Chaîne archivistique. De même, comme l'indique le rapport final de l'équipe transversale Description, le principe sous-jacent à ses travaux « est que des métadonnées détaillées et dignes de confiance sont essentielles pour garantir la production de documents fiables et la conservation de documents authentiques, et autres objets, dans les systèmes électroniques... [ce qui] plaide en faveur d'un régime de gestion des métadonnées « de bout en bout » qui définisse quelles métadonnées doivent être conservées et/ou transmises dans le temps, à quelles fins, par qui et comment elles doivent être conservées et validées⁹³. » Dans ce but, l'équipe Description, en collaboration avec l'équipe Modélisation, a élaboré un *modèle de spécification des métadonnées*. Ce modèle, très proche du modèle de référence pour un système ouvert d'archivage de l'information OAIS⁹⁴, est destiné à être utilisé, conjointement avec le modèle de Chaîne archivistique, comme point de départ pour la définition des spécifications d'outils automatisés qui pourront contribuer à la production, la capture, la gestion et la conservation des métadonnées essentielles pour les archives courantes, intermédiaires et définitives, ainsi qu'à l'identification des métadonnées devant être créées manuellement et de celles pouvant être synthétisées et/ou abandonnées à tel ou tel stade.

Vers des principes directeurs pour conserver des documents d'archives numériques authentiques

Les travaux d'InterPARES 2 ont mis en lumière l'absence d'outils pour assurer la conservation des documents numériques. Ce constat, pour le moins alarmant, concerne toutes les activités étudiées – artistiques, scientifiques et administratives – et touche aussi bien les producteurs individuels que les petits groupes, les projets subventionnés, les administrations publiques et les grandes entreprises. Chaque unité de recherche InterPARES – équipes sectorielles, équipes thématiques et équipes transversales – a identifié des politiques, des stratégies, des principes et des principes directeurs qui pourraient aider les producteurs et les services d'archives pendant la transition d'un archivage

⁹³ Description Cross-domain Task Force Report, 291.

⁹⁴ Le modèle OAIS décrit les rôles, responsabilités et fonctions d'un entrepôt numérique, ingestion, stockage « archivistique », gestion des données, planification de la conservation, administration et accès. Voir : <http://www.public.ccsds.org/publications/archives/650x0b1.pdf>. et pour la traduction française (2005) de la version 1.0 de la norme OAIS : http://pin.association-aristote.fr/lib/exe/fetch.php/public/documents/norme_oais_version_francaise.pdf.

analogique à un environnement d'archivage numérique. Le thème 3 a été chargé de rédiger les *Principes directeurs à l'usage des services d'archives en charge de la conservation définitive*⁹⁵, document destiné à accompagner les *Principes à l'usage des services d'archives*⁹⁶ élaborés par l'équipe transversale Stratégie et orientations. Parce que les études de cas se sont concentrées sur les producteurs, la collecte de « meilleures pratiques » de conservation a été maigre. Les Principes directeurs s'inscrivent donc dans une double perspective :

- Des actions qui devraient être conduites pour éviter certaines des situations observées dans les études de cas les plus problématiques
- Des actions qui devraient être conduites pour répondre aux préoccupations d'évaluation et de conservation identifiées au cours des travaux d'InterPARES 2.

⁹⁵ Voir l'annexe 21.

⁹⁶ Voir le *Policy Framework* en Annexe 19.