
Informatique 40 ans après: pour une typologie des mémoires à préserver

R.-Ferdinand Poswick

*Maison des Écritures, Informatique & Bible, asbl
11 rue de Maredsous, B-5537 Denée Belgique
cib@cibmaredsous.be*

RÉSUMÉ. Sur base d'une expérience de plus de 40 ans (1969-2012) dans le champs d'une informatique appliquée aux textes et documents, on tentera de donner ce qui nous semble la liste la plus complète des différentes traces d'une activité de ce type. Après quelques exemples montrant la source de cohérence d'une collection liée de près à la prise de conscience du passage de l'écriture alphaphonétique à l'écriture électronique, on décrira 20 ensembles en se demandant ce qui peut ou doit être conservé une fois la "mémoire vive" des acteurs disparue :

1. Lieux et locaux; 2. Personnes; 3. Analyses de projets et analyses informatiques; 4. Matériels; 5. Logiciels (progiciels) créés ou utilisés; 6. Journaux d'exécution; 7 Fichiers de données ou de programmes; 8. Données électroniques; 9. États imprimés; 10. Documentation administrative et courriers; 11. Les institutions; 12. Documentation technique et scientifique; 13. Éphémères techniques ou commerciaux; 14. Imprimés résultants de traitement informatisés; 15. Archives audio-visuelles; 16. Dossier de presse; 17. Réflexions et mémoires rédigés par les acteurs; 18. Publications techniques, scientifiques et synthétiques sur les applications; 19. Traces de télécommunications et courriels; 20. Inventaires électroniques.

ABSTRACT. On the basis of 40 years of computers' application in the field of texts and documents, this paper presents first the core of a coherent collection linked to the emergence of the consciousness of transition from alphaphotenic writing to electronic writing, and, then, what could be the complete list of memories to be kept when the living actors of this experience are gone!

20 fields of memorial will be described, with questions about « how to preserve what ? ».

MOTS-CLÉS : mémoires; traces; souvenirs; virtualisation; expérience humaine; transmission; informatique, écritures

KEYWORDS: memories, traces; remembrance; virtualization; human expertise; transmission; computer science; writings; scriptures

1. Introduction

Parce que les espaces de stockage n'étaient pas trop limités et que l'activité n'avait rien du caractère expéditif de travaux purement commerciaux, les équipes que j'ai dirigé depuis 1968 ont conservé des traces très diverses de la découverte progressive de l'écriture électronique – une découverte qui progressait également au rythme et au gré des évolutions technologiques.

Rien d'original, a priori, dans les travaux réalisés et dans les traces qu'on en a conservé? Il est évident que d'autres que les équipes du Centre Informatique et Bible à l'abbaye de Maredsous (Belgique), ont commencé, dans les mêmes années – et même avant – à utiliser des processus faisant appel, pour des segments de réalisations, aux machines électroniques programmables. Mais a-t-on gardé beaucoup de traces de la façon dont ils ont utilisé ces « calculatrices » en les transformant en machines pour les productions de base de la civilisation de l'écrit alphabétique, le livre (qui, au fil des ans, deviendra base de données en écriture électronique)?

2. Cohérence d'une Collection

La cohérence de la modeste « collection » qu'Informatique et Bible (I&B) cherche, aujourd'hui, à intégrer à un espace de mémoire plus large qui porterait sur l'informatique pionnière en Belgique, tient à une prise de conscience culturelle fondamentale induite par la nature-même des principales réalisations d'I&B. Cette prise de conscience est celle du passage de la civilisation fondée sur l'écriture alphabétique à une civilisation fondée sur l'écriture électronique.

En effet, le travail d'I&B a été originellement et principalement centré sur le texte, et, plus particulièrement, le texte de la Bible. Et non pas le texte pour en faire l'exégèse, la traduction ou un commentaire, mais avant tout pour l'éditer dans la tradition gutenbergiennne. Ce qui suppose une confrontation avec cette matière culturelle spécifique – il y faut aussi, bien sûr, l'exégèse, la traduction, le commentaire, etc – qui exige de maîtriser également le matériau textuel jusqu'au niveau typographique.

Or, on se trouve là, au lieu précis où l'on sera confronté à l'histoire de l'écriture, et, plus largement, à l'histoire de la communication humaine.

3. De l'écriture alphabétique à l'écriture électronique

La naissance de la Bible coïncide avec l'implantation du système de fixation des mémoires humaines sous forme de glyphes phonétiques plutôt que de représentations pictographiques. Toutes nos écritures alphabétiques dérivent de l'écriture proto-sémitique orientale qui s'est développée à partir de 1.500 avant notre ère à Babylone et qui se répandra en se diversifiant dans tout le Proche-Orient, puis

autour de la Méditerranée, en donnant naissance aux écritures grecques, coptes, phéniciennes, latines, arméniennes – et, plus tard, arabes, puis cyrilliques.

Cette découverte technologique de l'écriture alphabétique, un saut d'abstraction pour le cerveau humain¹, va faciliter l'échange commercial en créant le « contrat écrit », qui se traduira, au plan de l'organisation sociale en une fixation de règles ou de lois (la Thora) faciles à transmettre de générations en générations (l'Histoire) et à actualiser (le Commentaire, la Prophétie).

En travaillant à l'édition de ces textes et en adoptant, pour ce faire, les technologies de l'informatique naissante (le travail proprement informatique commence pour I&B en 1971), I&B est amené en permanence, et depuis plus de 40 ans, à confronter la culture qui est l'objet de son travail, aux outils de production et de communication de la culture nouvelle émergeant de l'utilisation de ce que nous finirons par nommer l'écriture électronique.

4. La typographie électronique

Pour la prise de conscience de ce basculement culturel qui est au cœur de la collection mémorielle d'I&B, l'obligation de commencer le parcours informatique par le métier de la typographie électronique, a permis de relativiser fortement ou de mieux situer la place exacte de l'« invention » de l'imprimerie dans l'évolution qui mène de l'écriture alphabétique à l'écriture électronique. Gutenberg et ses émules ont permis, grâce à la conception et à la mise en œuvre des « caractères mobiles » – une invention qui n'aurait pas été applicable à des systèmes d'écriture qui comptent de trop nombreux glyphes, comme l'écriture chinoise – d'accélérer le processus de production et donc de diffusion des mémorisations alphabétiques. Du coup, toute la chaîne culturelle de la lettre alphabétique s'est développée à une vitesse et avec une ampleur inconnues jusque là: après la Renaissance et le siècle des Lumières, premiers signes d'accélération de la machine culturelle, l'expansion de l'imprimé fait naître la culture technicienne, scientifique et industrielle qui finira par produire le calculateur programmé qu'est l'ordinateur.

Comment donc garder les traces de cette prise de conscience?

Une série de « premières », tant du point de vue des acteurs de ces travaux que du point de vue de ces travaux eux-mêmes (même limités à leur domaine restreint, mais si fondamental!), a permis cette prise de conscience d'un bouleversement culturel. Si l'on a conservé des traces, sera-t-il possible de les transmettre de façon qu'elles puissent aider les générations futures à mieux mesurer la continuité et les ruptures?

Quelques exemples permettront de mieux saisir ces prises de conscience et la façon dont on en garde des traces.

Ensuite, et de façon un peu plus systématique, on proposera et on évaluera une série d'éléments constitutifs de ces expériences en se demandant ce que l'on peut et

¹ Voir: Régis Debray, *Dieu, un itinéraire*, Odile Jacob, 2001.

ce que l'on doit en conserver, et comment.

5. Trois exemples

5.1. Informatiser l'orthographe

En 1970, on commençait l'installation des premiers IBM-360 à la Caisse Générale d'Épargne et de Retraite (C.G.E.R.- A.S.L.K.) de Bruxelles. Il n'y avait pas, à cette époque, en Belgique, d'École spécialisée en informatique², seulement quelques cours sur les nouveaux moyens de calcul donnés aux ingénieurs dans les universités et Grandes Écoles. On apprenait l'informatique « chez le constructeur ». L'avis donné par le conseiller scientifique d'IBM-Belgium aux deux moines bénédictins qui prétendaient traiter l'édition d'une concordance verbale et thématique de la Bible à l'aide d'ordinateurs, fut: « Devenez analystes et programmeurs pour pouvoir adapter ces techniques à la matière propre que vous voulez traiter et qui est bien éloignée des préoccupations et objectifs de la majorité de ceux qui utilisent ces machines à calculer en physique nucléaire, statistique et autres, où les mathématiques sont reines ». La femme de M. David Hirschberg, Lydie, s'intéressait aussi à ces nouveaux outils; elle avait fondé le premier laboratoire de linguistique assistée par l'informatique à l'Université Libre de Bruxelles (U.L.B.) et avait développé un programme de terminologie multilingue (5 langues européennes) sous le nom de DICAUTOM³. C'est elle qui orientera le choix des apprentis informaticiens vers le COBOL qui, au-delà de l'ASSEMBLER absolument indispensable à cette époque pour des routines un peu pointues, était mieux adapté au traitement des caractères par rapport au FORTRAN qui traitait mieux les chiffres!

Et voilà nos deux bénédictins engagés dans les cours intensifs qui formaient analystes et programmeurs chez IBM-Belgium, tout en lançant presque simultanément le travail de mise sur cartes perforées de 80 colonnes des 300.000 fiches de format bibliographique créées manuellement en vue de la constitution de cette table analytique de la Bible qui paraîtra sous forme imprimée en février 1974⁴.

Première confrontation de la lettre alphabétique avec l'écriture électronique: à l'époque (1971), la filière d'utilisation des IBM-360 n'offrait pas, avec son code EBCDIC, la possibilité d'adresser directement des lettres accentuées (nombreuses et significatives en français!). Il nous fallait, en plus, pouvoir signaler dès le départ, les lettres capitales qui n'étaient pas appelées par leur position (après une ponctuation forte, par exemple). Il a donc été nécessaire de prévoir et programmer, puis faire encoder les données, avec un interface qui nous permette, en fin de traitement, d'obtenir une typographie française conforme à l'orthographe. Il était donc nécessaire de « coder » les mots français: majuscule (hors position) = « / »; accent aigu = 1; accent grave = 2; accent circonflexe = 3; tréma = 4; cédille = 5, en veillant

² L'Institut d'Informatique des FUNDP (Namur) s'ouvre en 1971.

³ Ce qui deviendra le noyau de la Banque de données terminologique Eurodicautom encore utilisée par les 9.000 traducteurs des institutions européennes.

⁴ *Table Pastorale de la Bible*, Paris, Lethielleux, 1974, 2011².

à ce que ce codage soit très mnémotechnique pour ceux qui perforaient les données⁵. Au programmeur de gérer la séquence des codes pour rétablir, en sortie, la graphie correcte.

Par la suite, il s'agira de faire le même de genre de sport pour représenter tous les accidents grammaticaux et graphiques de l'hébreu/araméen, du grec, de l'arabe, du syriaque... puis de plusieurs langues modernes autres que l'anglais! Et cela jusqu'au jour où des glyphes de différents jeux de caractères seront offerts sous forme d'images adressables, d'abord de façons variées en ASCII, puis de façon théoriquement unifiée à travers l'UNICODE.

Et si cela se programmait à l'origine pour des photocomposeuses électroniques (Hell-Digiset-Siemens de l'Imprimerie Nationale à Paris, à partir de 1972) ou des imprimantes à laser (IBM-3800, à partir de 1979, les premières à permettre d'imprimer des caractères hébreux à peu près corrects), ce fut le même effort pour afficher de tels caractères sur les premiers écrans graphiques TEKTRONIX (1980).

5.2. Du livre à la base de données

Si tout le texte de la Bible hébraïque vocalisée et cantilée fut encore enregistré, entre 1976 et 1980, dans le type de codage évoqué plus haut sur des cartes perforées de 80 colonnes (après avoir été entièrement transcrit sous cette forme sur des bordereaux de 80 colonnes et 25 lignes), il fallut attendre l'avènement des micro-ordinateurs, et notamment le premier IBM-XT (1984) auquel il fallait adjoindre un disque dur extérieur avec suffisamment de mémoire (60Mb), pour présenter la première Base de données de l'hébreu massorétique associée à un programme de recherche multicritère⁶. Cela se passait en 1984-1985 et fut présenté à la première conférence organisée par l'Association Internationale Bible et Informatique (A.I.B.I., fondée par I&B en 1982) à Louvain-la-Neuve en 1985, puis à la Conférence annuelle du groupe de contact CARG (Computer Assisted Research Group) récemment créé au SBL-meeting à Chicago la même année, avec notre associé Ronny Benun de Brooklyn.

Ce qui, jusqu'alors, s'était toujours présenté sous forme d'un « livre » (les Concordances de la Bible, le seul livre jugé digne, jusqu'à très récemment, de ce genre de travail démentiel depuis qu'Hugues de Saint-Cher en avait fait un premier modèle vers 1235 au couvent dominicain de Saint-Jacques à Paris, une réalisation qui exigeait un très long travail de préparation, de fabrication, puis de distribution d'un bloc de papier qui pesait de 2 à 4 kg, voire plus!), cela devenait soudain un autre objet, dont les éléments se stockaient sur des bandes magnétiques, puis sur 135 disquettes de 5¼ ou un disque de 60Mb, accessibles par des programmes écrits en BASIC.

⁵ Exemple: *été* = 1ET1E; *père* = P2ERE; *Père* = /P2ERE; *êtes* = 3ETES; *aiguë* = AIGU4E; *ça* = 5C2A; etc.

⁶ *Mikrah-Compucord*, Maredsous-Brooklyn, 1985.

5.3. Livre et Minitel

Le 30 juin 2012, Orange a arrêté définitivement ce qui restait des services Teletel. Les terminaux Minitel ont commencé à être distribués largement par France-Télécom au cours des années 1984-1985, suite aux essais réalisés en Ile-et-Vilaine en 1982 et à l'ouverture, cette même année, de l'Annuaire électronique.

Dès 1987, pour la première fois dans le monde de l'édition, Informatique & Bible présentait à Paris, simultanément, la version imprimée du *Dictionnaire Encyclopédique de la Bible* et sa version en Base de données accessible avec le programme de recherche STAIRS sur le serveur GENETEL (Bordeaux). Et, l'année suivante (1988), avec cette même application, I&B établira la première liaison d'un Minitel entre la France et Israël à l'occasion de la 2^{ème} Conférence de l'Association Internationale Bible et Informatique (A.I.B.I.)⁷. Ce même dictionnaire dans sa 3^{ème} édition est désormais accessible, à l'aide du moteur de recherche *Knowhowsphere* pour Bases de données enrichies, tant dans le WEB que dans une version pour mobiles (*KHSmobi*), avec des bibliographies mises à jour jusque fin 2010.

6. Que conserver d'expériences pionnières dans notre champ d'application et comment?

Archiver c'est garder, pour le moyen ou le long terme, des traces d'une réalité qui a été vécue ponctuellement ou de façon récurrente (avec des variantes) durant une période plus ou moins longue dans un environnement qui ne cesse de se modifier au fil du temps.

Quelles sont les traces dont il faut garder mémoire? et sous quelle forme conserver cette mémoire? Je donne ici ce qui me semble constituer la liste des différentes traces d'une activité. Par facilité et par souci d'être concret ce descriptif est fait sur base des applications pionnières de l'informatique à la Bible, aux textes, aux documents, aux archives durant une période de 40 ans s'étendant de 1969 à 2009, par l'équipe d'Informatique & Bible. I&B conserve de nombreuses traces matérielles ou virtuelles pour chacun des éléments évoqués.

Quelle serait les différences entre un tel ensemble que l'on peut considérer comme une "mémoire vive" et ce qui en serait conservé hors la présence des acteurs?

Voici les éléments de cette « mémoire vive » dont on trouvera le détail discuté dans l'Annexe 1.

⁷ R.-F. Poswick et J. Bajard, *COMPUCORD de Mikrah: une base de données en accès interactif sur micro-ordinateur pour tout le texte massorétique (B.H.S.) et ses équivalents lexicaux dans la Bible anglaise (R.S.V.)*. Actes du Second Colloque International Bible et Informatique: Méthodes, Outils, Résultats, Jérusalem , 9-13 juin 1988, Paris-Genève, Champion-Slatkine, 1989, pp. 87-98.

- Les *lieux* et les *locaux* : ils peuvent avoir une signification culturelle, mais ils peuvent aussi conditionner le travail et en garder des traces.

- Les personnes: leur cheminement individuel et leur interaction dans une équipe est à la base de tous les projets.

- Les *analyses de projets* et les *analyses informatiques* fonctionnelles: ce sont les cadres qui permettent de comprendre les programmations.

- Les *matériels*: si on peut les conserver, ils sont souvent l'ossature principale d'une présentation muséale. Mais il est difficile et coûteux de vouloir les garder opérationnels.

- Les *logiciels*: ils peuvent être créés ou simplement utilisés; ils peuvent être une intégration de routines créées ou réutilisées. La connaissance des langages de programmation est nécessaire à leur conservation.

- Les *journaux d'exécution* signalent les accidents de parcours, les étapes de recherche, les essais, les durées de réalisation.

- Les *fichiers de données ou de programmes*: s'ils sont conservés, en a-t-on le catalogue? Peut-on distinguer ceux qui sont opérationnels de ceux qui sont conservés à titre purement archivistique?

- Les *données électroniques*: que peut-on savoir de l'intérêt que pourrait représenter, dans le futur, certains éléments qui ont été conservés sous forme électronique mais qui demandent un certain travail pour une exploitation actualisée (fichiers en code EBCDIC, par exemple)?

- Les *listings*: il y en a de moins en moins, mais toute la période pionnière de l'informatique en a produit énormément. Conserver l'idée des masses que cela constitue a-t-il un intérêt?

- La *documentation administrative et comptable*: elle donne des dates précises et situe les projets dans l'économie et la société.

- Les *institutions*: elles sont le cadre dans lequel les projets ont été réalisés. Assemblées générales, Conseils d'Administration, Conseils de Direction reflètent et conditionnent les développements.

- La *documentation technique et scientifique*: elle va du manuel d'installation d'un matériel à une bibliothèque spécialisée dans les domaines utiles aux applications.

- Des « *éphémères* », s'ils ont été conservés, permettent de comprendre l'offre globale de l'informatique dans les domaines d'application à différents moments des réalisations.

- Des *livres imprimés* ou des outils de distribution de bases de données électroniques, représentent, dans nos domaines, des traces de résultats.

- Les *archives audio-visuelles* font revivre lieux, personnes et matériels en activité. Mais c'est souvent très épisodique et peu technique.

- Un *dossier de presse* montre l'écho des réalisations pour un plus large public.

- La *réflexion synthétique* sur l'activité ou les réalisations est souvent représentée par des articles dans des périodiques ou des Actes de colloques et conférences.

- Il peut exister également des *publications techniques* sur les applications.
- Des *bases de données* donnant de l'information sur différents aspects des applications peuvent exister. Sont-elles figées ou dynamiques?
- Dans cette ligne un *Inventaire électronique* de toutes les sources diverses de mémoire peut également exister et permettre de naviguer et de rechercher ce qui semble intéressant. Comment conserve-t-on de façon dynamique ce genre de mémoire?

7. Conclusion:

Si l'on tente de regrouper les différentes sources qui forment l'état 'bilantaire' d'une entité historique ayant travaillé pendant près de 40 ans dans le domaine de l'informatique appliquée, on peut, me semble-t-il, délimiter certains ensembles dont la permanence, sans un traitement conservatoire volontaire, sera plus ou moins rapidement mise en question:

a) L'expérience vivante et la mémoire biologique des acteurs pourraient constituer les traces qui disparaîtront le plus vite et souvent le plus complètement. Il y a, cependant, quelques traces que l'écriture électronique peut aider à conserver: images fixes ou mobiles, sons d'interviews, dossiers personnels des différents collaborateurs.

b) Les lieux et locaux subsisteront plus longtemps, mais ils peuvent ne plus garder aucunement les configurations qui ont servi de cadre à l'activité. L'image électronique peut suppléer dans une certaine mesure.

c) Les créations complexes dans lesquelles données et algorithmes sont mêlés pour représenter un ensemble interactif aléatoire, éventuellement susceptible d'intégrer, à la volée, de nouveaux éléments, risquent d'être liées à la permanence des acteurs qui les ont mises en œuvre, qui les mettent à jour et les animent.

d) Tout ce qui peut représenter un 'écrit' imprimable (la très grande majorité des éléments décrits ci-dessus), peut subsister, soit tel quel si les conditions de conservation (espace, humidité) sont favorables et les espaces disponibles, soit sous forme d'un fichier électronique de textes ou d'images. On notera au passage que bien des 'impressions' des premières époques d'imprimantes (poussière d'encre) ou de télécommunications (fax à papier thermique) auront probablement disparu si elles n'ont pas été recapturées électroniquement.

Mais, sans un type d'inventaire ou de descriptif qui relie ces ensembles de documents, ils peuvent vite devenir sans intérêt car on risque de ne plus être en mesure d'en percevoir l'intérêt. Le poids et l'encombrement poussent à l'élimination de ces données qui disparaissent à tout jamais si elles n'ont pas fait l'objet d'une virtualisation au moins sélective!

e) Il y aurait donc à privilégier la conservation des sources synthétiques (Annexe 1, n° 17 et 18) si l'on n'a pas la possibilité de créer une Base de données structurée reliant l'ensemble des traces de mémoire ... et peut-être de tirer une

impression sur papier non-acide ou sur microfilm des éléments les plus importants à mémoriser.

f) Et, dans tous les cas de mémorisation sous forme d'écriture électronique, les différentes normes prudentielles de conservation devraient être d'application. Sur ce sujet, on verra, en Annexe 2, les remarques judicieuses de ma collaboratrice Yolande Juste à l'occasion de la constitution, le 29 octobre 2009, du Réseau de Préservation des Patrimoines Informatiques.

Annexe 1: R.-F. Poswick, Vingt éléments de mémoire conservés chez I&B

1. **Les lieux ou locaux:** ils se sont succédés divers; ils ont, ensuite, été utilisés autrement et ont été transformés pour la plupart. Les seules traces peuvent être des descriptions, des photographies et/ou des plans. Un lieu historique peut devenir un cadre muséal. Le mode d'implantation de certains matériels et notamment l'installation d'un câblage spécifique sont déjà souvent liés à une analyse de configuration de l'environnement de travail. L'évolution des types de câblage et de la nature des connexions dit quelque chose de l'utilisation des matériels électroniques à une époque donnée. Conserve-t-on, et comment, une trace de ces choix? Conserve-t-on des exemples de types de matériels de câblage?

2. **Les personnes:** acteurs ou collaborateurs qui se se sont succédés nombreux (plus de 150 personnes). Les traces de leur mémoire pourraient être collectées sous forme d'interview⁸. On dispose de photographies documentées et de documents ou dossiers qui peuvent être attribués au travail spécifique de telle ou telle personne. Un dossier administratif⁹ et social existe également sur chaque collaboration. Pour ceux qui ont publié sur leurs travaux ou sur d'autres domaines, une bibliographie peut être établie. Une notice biographique peut rassembler beaucoup de ces éléments.

3. **Analyses de projets et analyses informatiques fonctionnelles:** Elles existent dans deux lignes différentes, soit comme analyse synthétique des données et des conditions de faisabilité d'un projet, soit comme analyse des méthodes informatiques à utiliser. Beaucoup sont sous forme de manuscrits ou d'imprimés; il y a quelques fichiers électroniques (surtout depuis l'existence des 'traitements de texte' sur micro-ordinateurs). Ces dossiers font partie des Archives conservées.

4. **Les matériels** utilisés et en fonction desquels les analyses ont été créées: du travail initial sur les IBM-360 (et évolutions subséquentes) à la CGER (Caisse Générale d'Épargne et de Retraite à Bruxelles), il reste des bandes magnétiques (figuratives, car toutes les données ont migré vers des supports actualisés) et des machines à perforer – d'autres matériels de ce type ont été récupérés a posteriori et ont été conservés à titre muséal sans plus fonctionner. Différents exemplaires d'autres machines gardent mémoire de différentes étapes de travaux: on a conservé, presque toujours, toute la documentation qui permettait leur utilisation. Des photographies de matériels sur leurs sites d'utilisation existent également. Ce genre de matériels est souvent l'ossature principale d'un espace muséal. La question des modes de préservation et de présentation est abondamment traitée par ailleurs au cours de cette conférence.

⁸ Le projet de la Faculté d'Informatique des FUNDP à Namur tel que proposé par Marie Gevers, est un bon exemple dans ce sens (voir: M. d'Udekem-Gevers, *La Machine mathématique IRSIA-FNRS (1946-1962)*, Bruxelles, Académie royale de Belgique, 2011, pp. 169ss).

⁹ Notamment des statistiques sur la durée des différentes phases de travaux d'après les relevés de Time-Sheets de chaque collaborateur, et le Rapport journalier que chacun produisait.

5. **Les programmes créés ou utilisés**, à condition qu'ils soient documentés. Ils sont conservés sous forme de listes imprimées (ou, pour les plus anciens, sur cartes perforées), et/ou sous forme de fichiers électroniques. Il s'agit ici de "sources" créées en vue d'une application électronique spécifique ou, selon le jargon du métier, 'propriétaire'. Mais il existe également une bibliothèque bien archivée d'applications logicielles extérieures utilisées (Mini, PC et autres), avec leur documentation. Outre un inventaire de ces différents produits logiciels, une description de l'usage qui en fut fait devrait compléter le dossier.

6. **Les journaux d'exécution** quand ils ont été minutieusement tenus sous forme manuscrite (cahiers d'exécution), ont été classés à part dans une première période, puis, ensuite, classés avec les différentes applications ou projets. Certains ensembles peuvent avoir été virtualisés sous forme de photographies numériques de l'ensemble d'un dossier, photographies accessibles dans l'Inventaire général des traces de mémoire.

7. **Les fichiers de données ou de programmes**: un inventaire électronique de tous les fichiers créés et conservés existe. Il s'agit souvent du dernier état d'un programme, du dernier état de l'élaboration des données pour une Base de données, du dernier état électronique avant une photocomposition typographique.

8. **Les données électroniques** de toutes les applications depuis le début des travaux. Certaines de ces données, indépendantes des programmes et machines qui les ont créées peuvent demander une nouvelle programmation pour être réellement utilisables. Ce qui suppose des migrations tant au niveau des supports qu'au niveau de certaines façons de représenter les données (ex: du code EBCDIC au code ASCII, des codages sur plusieurs caractères en ASCII, comme pour l'hébreu, vers des glyphes représentés par un code UNICODE, etc). Les données les plus récentes (10 dernières années) en format XML (un formatage de données fixé dans les années 1990ss), migrent déjà plus facilement, sous réserve de documentation valable sur les règles de codage.

9. **Les listes (listings)** états imprimés des données traitées et de beaucoup de programmes existent pour différentes étapes de certains travaux ou seulement pour l'état final. Ces listes sont conservées dans des caisses qui représentent un volume important. Si l'on ne peut tout conserver, il ne semble pas suffisant de ne conserver qu'une ou deux pages de chaque 'sortie' d'imprimante; il semble intéressant de conserver une idée des 'masses' de listing que représentent les résultats de travaux importants pour les époques où l'on imprimait pratiquement tous les résultats. Pour les époques plus récentes, des échantillons de résultat sous forme imprimée se trouvent dans les dossiers d'exécution; mais les données complètes sont conservées sous forme de données électroniques.

10. **La documentation administrative et comptable** relative aux différents travaux est conservée dans des boîtes d'archives selon deux classements: a) des documents variés regroupés autour d'un projet; b) un répertoire chronologique de toutes les pièces produites depuis 1980 (CHRONO).

11. **Les institutions**: Il existe également des comptes-rendus des Conseils de Direction, Conseils d'Administration, Assemblées Générales des différents cadres sociaux (asbl, S.A., etc) dans lesquels se sont déroulés les travaux. Une virtualisation ou numérisation de ces dossiers est possible.

12. Une **documentation technique et scientifique** a servi de référence à tous ces travaux. On y trouve: des manuels d'utilisation de matériels et de logiciels, des livres de réflexion sur l'informatique et les disciplines connexes intéressants les travaux, de la documentation imprimée sur un ensemble d'applications spécifiques (littérature, linguistique, langues, langages de programmation, intelligence artificielle, archivistique, informatique documentaire, bibliothéconomie, gestion de projets informatiques, etc). Cette bibliothèque, complétée par une bibliothèque de périodiques, est en bonne partie cataloguée sous forme électronique.

13. Il existe une collection d'“éphémères”: prospectus, descriptifs, articles, gadgets publicitaires, y compris une collection de “sacs publicitaires”; cela peut être inventorié et catalogué électroniquement.

14. Une série d'imprimés (livres) représente des résultats publiés de travaux. Une bibliographie historique peut en être constituée.

15. **Archives audio-visuelles:** photographies (pellicules et albums classés), bandes vidéos, films avec interviews ou autres de différents acteurs. Ce matériau peut être enregistré sous forme numérique.

16. Un **dossier de presse** rassemblant des coupures de journaux et revues a été tenu à jour. C'est un ensemble qui peut aussi être 'numérisé'.

17. Il peut exister, en outre, une série de productions qui représentent une **réflexion ou mémoire synthétique** des travaux réalisés ou, plus largement, une confrontation analytique ou académique avec les milieux qui développaient et appliquaient l'informatique dans des domaines connexes, parallèles ou concurrents, notamment: contributions à des colloques et conférences. Une bibliographie documentée peut reprendre l'information sur ces publications.

18. Il existe également des **publications techniques, scientifiques et synthétiques sur les applications.**

19. Dans le monde des données offertes en **télécommunication**, il existe des ensembles de données, structurées ou non sous forme de bases de données interactives.

20. Un **inventaire électronique** de l'ensemble des traces décrites ci-dessus peut exister ou être envisagé.

Annexe 2: Yolande Juste, *Méthodologie intellectuelle, matérielle et logicielle, Informatique & Patrimoines, Séminaire organisé par la Maison de la Métallurgie et de l'Industrie de Liège et la Maison des Écritures, Informatique & Bible, Maredsous, 29 octobre 2009*

Plusieurs institutions collectent et mettent à disposition à travers Internet des "archives" plus ou moins considérables. (jusqu'à 3 petaoctets pour archive.org) de textes, de sons, vidéos et même de logiciels. Ces "archives" semblent être plus de l'ordre du stockage accumulatif de données diverses en leur état au moment de leur entrée dans ces archives. Faut-il se contenter de ce type d'accumulations ou doit-on travailler à un autre mode de conservation à moyen et long terme dans le secteur des patrimoines informatiques?

– **QUI:** la préservation d'un patrimoine ne doit pas être réservée aux spécialistes de l'archivage mais devrait concerner chaque créateur de données, y compris le particulier à qui l'on impose par exemple les documents administratifs électroniques et qui est donc responsable de leur conservation à plus ou moins long terme suivant des règles légales.

– **QUE, QUOI:** étant donné la croissance exponentielle du nombre d'informations électroniques créées chaque jour par chacun, la première démarche consiste à définir CE qu'il faut conserver, le type de données, et ceci à chaque niveau de création, du particulier, aux administrations en passant par les PME, PMI.

– **DONT:** quels états des données faut-il conserver? Si on tente, par facilité, de tout conserver, on atteindra vite des volumes de données non maîtrisables, même si

les capacités de mémoire ne cessent d'augmenter avec des encombrements toujours moindres, mais avec des coûts de gestion importants.

– **OÙ** faut-il conserver ces données? faut-il se fier à des mémorisations décentralisées sur quelques gros serveurs détenus par des sociétés privées? Quelles obligations de conservation ont ces sociétés au-delà du minimum légal? Sait-on toujours où se trouvent ces données, dans quel pays? quel continent? Et quid en cas de défaillance ou de faillite d'un tel opérateur?

– **COMMENT** conserver ce qui doit l'être? Il faut distinguer plusieurs types de données:

- Pour les *données "imprimables"* (textes, photos) une norme d'archivage, PDF/A (ISO 19005) tente de s'imposer. Si l'idée de base est bonne: indépendance, autonomie..., elle impose certaines restrictions (non inclusion d'objet dynamique, interdiction du lancement de code script ou exécutable, inclusion de toutes les polices de caractères, UNICODE, interdiction de chiffrement et de sécurité...) et son application est questionnable. Si chacun connaît et produit des PDF, combien de ces fichiers sont-ils compatibles avec la norme PDF/A? Expérience faite, presque aucun document PDF n'est directement compatible PDF/A et certains PDF/A ne passent même pas le test de validation avec succès. De plus cette norme ne s'applique aujourd'hui qu'à des documents "imprimables", donc des textes et des images. Sont exclues les données conservées en "base de données" si elles ne sont pas exportées sous formes d'états imprimés, les sons, les vidéos....

L'archivage PDF/A est donc une démarche volontaire, complexe et coûteuse dans la mesure où les logiciels de validation proposés sont eux-mêmes coûteux. Olaf Drümmer affirme avec un optimisme certain, que cette norme s'imposera le jour où la masse critique sera atteinte, ce qu'il évalue à 2010!

- Les *données non imprimables*, ou non imprimées (fichiers de données, bases de données...), doivent être conservées dans un format indépendant de tout logiciel. Le format aujourd'hui reconnu est l'XML en UNICODE, accompagné de sa structure (DTD ou schémas) et de sa mise en forme (CSS...).

- Les *données interactives*? tels les sites WEB? aucune directive ne les concerne. Une initiative américaine *archive.org* archive TOUTES les versions de TOUS les sites web. Mais là encore, la démarche est questionnable de plusieurs points de vue, aussi bien des droits d'auteur (chacun a le droit d'exclure son site de cette procédure robotisée, par une démarche opt-out!) que de la pérennité (quelles obligations dans la durée?).

- Les *logiciels*: Là encore la nuance s'impose: Les logiciels propriétaires, ceux que l'on a créé et dont on possède les "sources" en langage clair, lisible et donc "imprimable" peuvent être conservés mais pas indépendamment de l'environnement sur lequel ils "tournaient". Les programmes ou logiciels que l'on a simplement utilisés et dont on ne possède pas les sources pourraient être conservés avec leur environnement d'exécution.

La conservation de ces logiciels ne peut se concevoir qu'en lien avec le matériel sur lequel ils tournaient.

Au-delà de ces normes (PDF/A, XML...) il ne faut pas oublier la question des supports. Tout support électronique doit être vérifié périodiquement quant à sa fiabilité de lecture, le slogan "*avec P..... vous gravez pour l'éternité*" n'est que publicité mensongère. Seule la migration des supports garantit leur lisibilité future. Mais dans ce cas, seules les données sont conservées comme patrimoine. Les supports eux-mêmes ne peuvent être conservés QUE comme une trace matérielle de l'évolution de ces techniques, et dans ce cas, un seul spécimen suffit.

– **CEPENDANT**, l'idéal n'étant jamais atteint, la conservation, même non normalisée ni organisée, ne vaut-elle pas mieux que l'oubli? À charge aux historiens de faire revivre ces patrimoines!

– **NÉANMOINS**, chacun peut déjà apporter sa contribution à une meilleure conservation en appliquant des règles simples de nom de fichier, de copie de sécurité et en utilisant, au jour le jour, les outils les plus adaptés à une certaine indépendance et longévité (odt plutôt que doc, xml plutôt que mdb ou dba...).

– **TOUTEFOIS** la première question n'est-elle pas: "le monde virtuel qui nous envahit n'appellera-t-il pas une conservation virtuelle?". Ou encore: "comment adapter le principe de conservation au monde virtuel alors que la conception même de la conservation reste très matérielle?".

Ces réflexions constituent la modeste contribution d'informaticiens qui, depuis plus de 30 ans, sont passés des main-frames aux PC, des cartes perforées aux clés USB en passant par les disquettes de 8", 5" et 3", les bandes magnétiques, les disques optiques, les cd et les dvd.

Tout cela au prix de nombreuses migrations, mais sans aucune perte de données et en conservant toutes les données exploitables aujourd'hui, même si ce n'est pas dans leur environnement d'origine.