



Comment on préserve l'authenticité ?

- Pb
- Cher
- Transfert dans un nouvel environnement système : conversion de format / transfert dans un autre OS
- Copie simultanée à plusieurs endroits
- Technique visant à substituer à un élément de matériel informatique un logiciel
- Technique très employée pour le jeu vidéo
- Mise d'un UVC Universal Virtual Computer qui serait capable de lire tout fichier ou type de document
- Intérêt de cette stratégie est en cours de réévaluation par certains acteurs de la préservation comme une solution à moindre coût pour refaire certains fichiers obsolètes
- Principe : le document à pérenniser doit être conservé avec tout ce qui permettra de le lire et de l'interpréter dans le futur
- Encapsulation
- ASCI
- XML
- Choisir des formats de MD pérennes
- Tester et choisir des outils
- Elaborer des recommandations
- Prendre en compte les exigences de sécurité de SI
- Réaliser en mode "production"
- RETEX : Programme de Préservation numérique de BAC
- Connaître les limites
- Solutions techniques
- Suppression des arrière-pensées
- Pratiques, plans et politiques organisationnelles
- Humains (nombre d'ETP)
- Financiers
- Techniques (ressources de calcul notamment)
- Méthodologiques

Des engagements de service

- Exemple : SCAPE Policy éléments
- S'inspirer de modèles existants
- Démarche de gestion des risques
- Techniques
  - Rafraîchissement
  - Migration
  - Copie simultanée à plusieurs endroits
  - Réplication
  - Emulation
  - Encapsulation
  - Choisir des formats de MD pérennes (ASCI, XML)
- Tester et choisir des outils
- Elaborer des recommandations
- Prendre en compte les exigences de sécurité de SI
- Réaliser en mode "production"
- RETEX : Programme de Préservation numérique de BAC
- Connaître les limites
- Solutions techniques
- Suppression des arrière-pensées
- Pratiques, plans et politiques organisationnelles
- Humains (nombre d'ETP)
- Financiers
- Techniques (ressources de calcul notamment)
- Méthodologiques

ISO 5127:2017 "measures and actions applied to prevent, arrest or delay deterioration of a document or other material object"

- IP, trad, EDC : ensemble de principes, politiques, règles et stratégies destinés à prolonger l'existence d'un objet en le maintenant dans un état propre à l'utilisation, soit dans son format original soit dans un format plus pérenne, tout en gardant intacte la forme intellectuelle de l'objet.
- IP, trad, EDC (digital preservation) : Processus spécifique visant à maintenir les objets numériques à travers le temps et par-delà différentes générations de technologies, indépendamment du lieu de stockage.
- DAI, OASIS (long term preservation) : fait de maintenir sur le Long terme l'information
- Immédiatement compréhensible par la Communauté d'utilisateurs cible, avec les preuves de son authenticité.
- Entité "identification de la pérennisation" OASIS
  - Développer des modèles d'emballage et des plans de migration
  - Développer les stratégies et les normes de pérennisation
  - Veille technologique
  - Suivi de la communauté d'utilisateurs cible
  - Les différentes fonctions sont développées dans la norme et devront être prises en compte dans la modélisation
- Entité "administration" OASIS
  - cette entité assure les fonctions et services relatifs à l'exploitation d'ensemble du système d'archivage. Les fonctions de l'Entité « Administration » induisent : la proposition et la négociation des Protocoles de versement avec les Producteurs, la vérification des versements pour s'assurer de leur conformité aux normes d'archivage ainsi que la gestion de la configuration du matériel et des logiciels du système. Elle fournit aussi les moyens techniques pour contrôler et améliorer l'exploitation de l'Archive, ainsi que pour inventorier, rendre accessibles et migrer/mettre à jour les contenus de l'Archive. Elle est également responsable de l'entretien et du maintien des normes et politiques de pérennisation applicables à l'Archive, de l'aide à l'utilisateur et de l'actualisation des requêtes enregistrées.
- Restoration
  - ISO 5127
  - Restoration de données
  - IP, trad, EDC : Processus consistant à récupérer des données ou des données d'archives sous la forme de bits à partir d'un support numérique défectueux, endommagé, dégradé ou corrompu, et à effectuer différentes opérations pour restaurer l'intégrité des données ou documents récupérés.
  - Def. OASIS : période de temps suffisamment longue pour qu'il soit nécessaire de prendre en compte les changements technologiques qui ont un impact sur l'information détenue par un OASIS, et notamment la gestion des nouveaux supports et formats de données ainsi que l'évolution de la Communauté d'utilisateurs cible. Cette période n'est pas limitée dans le temps.

2 - Concepts archivistiques & diplomatiques

- Au regard de la norme OASIS
  - Distinction entité "identification de la pérennisation" (qui "pense"/ "administration" (qui "fait") = entité "stockage" pour la gestion des supports
- Au regard de la chaîne archivistique traditionnelle
  - Collecte, évaluation, sélection
  - Accès
  - Liens à renforcer par rapport à la situation actuelle, y compris sur sélection / évaluation ?
  - Inclusion de ces aspects dans le sujet "conservation" ?
  - Prévoir en mesure de statuer sur la capacité à prendre en charge un fichier
  - Prendre en compte les modes d'accès dans la politique de conservation, voire la politique de collecte
  - Interconnexion très étroite des différents sujets, davantage que dans le monde analogique
- Au regard des réflexions d'interPares
  - Constat : impossibilité de conserver en l'état un document d'archives numérique, il n'est possible que de préserver la capacité de reproduire ce document dans l'espace et le temps
  - Les 3 critères du document d'archives : Authenticité, Fiabilité, Exactitude
  - Pose de questions sur ce qu'est le document d'archives numérique et ce qu'implique conserver un tel document à travers le temps. Une appropriation de ces travaux est essentielle pour bâtir des stratégies de conservation sur des bases solides.
  - Approche très globale, notion de chaîne de Préservation (COP)
    - Prend en compte les différents systèmes dans lesquels va passer le document au cours de sa vie
    - Faut-il aller jusqu'à dans l'étude ?
- Au regard de la norme NF Z 42-013
  - La pérennité n'est pas la préoccupation principale de NF Z 42-013 mais le sujet est néanmoins traité et le SAE doit garantir cette pérennité
  - Il n'est donc pas possible d'opposer archivage intermédiaire (= vocation probatoire, intégrité) et archivage définitif (= pérennisation)

3 - Caractéristiques de l'objet à conserver

- Objet numérique est de nature complexe et dynamique
- Le document numérique = des composants numériques et des relations entre ces composants (interPares)
- Hétérogénéité des modes de production
- Hétérogénéité des endroits où se trouve l'information qualifiée et intéressante
- Croissance exponentielle de la masse
- Numerique natif / numérisé
- Préservation des logiciels est un défi à part entière
- Obsolescence des supports
- Obsolescence des formats
- Obsolescence des modes d'exploitation
- Obsolescence des usages

4 - Connaître & expertiser

- Développer une expertise
  - Disposer d'une base de ressources
    - Bibliographie
    - Normes & standards
    - Annuaire des acteurs
  - cf. livrables produits dans le cadre de FAMOA
  - Expertise archivistique appliquée au numérique
    - Structure de l'information numérique
      - Développer des outils et des registres pour évaluer la qualité des sets de métadonnées
      - Participer aux commissions de normalisation nationales et internationales
      - Projet MADRAS IP
    - Formats de métadonnées
      - maintenir plusieurs langages de développement, afin d'être en mesure de réutiliser les différents outils qui existent, il faut en effet comprendre comment les formats sont structurés, travailler à la réalisation des validateurs dédiés
      - avoir une appétence pour le domaine
      - être curieux et proactif, savoir faire de la veille, maintenir des outils
    - Formats de fichiers
      - RETEX CINES : se reposer sur les producteurs pour les formats très spécifiques ou les formats maison
      - RETEX CINES : expert format
      - RETEX CINES : faire de la veille
      - Missions de la chaîne de veille :
        - implémentation et maintien des différents outils
        - développement et maintien de FACILE
        - publication du code source du validateur sur une forge
        - assistance aux utilisateurs pour les fichiers KO (une aide peut être demandée en ligne)
        - participation à des groupes de travail
        - Rapports du Digital Preservation Coalition
      - Liens avec les communautés métier
      - Influencer la communauté métier
      - Ex : TEL
    - Technologies émergentes et innovantes
      - Veille sur les outils disponibles sur le marché
      - Test et parangonnage
    - Méthodologie
      - Veille sur les ressources identifiées
      - Participation aux travaux de recherche des groupes les plus actifs, notamment au niveau international
      - Lancement de travaux universitaires (programme de recherche)
      - Participation aux travaux de normalisation
      - Participation aux conférences et événements internationaux
      - Participation à la diffusion des connaissances
      - Former
      - Publier

5 - Agir

- Supports/stockage
  - Pilotage du stockage
    - Qui l'assure ?
    - au cours du cycle de vie ?
    - Transformer ou pas ?
    - contrôle des transformations ?
  - Migration support
    - RETEX INA montre que la question demeure d'actualité, au moins pour l'audio-visuel
    - RETEX INA : ensemble de règles et algorithmes permettant d'organiser l'information dans un objet numérique, par exemple : spécifier le codage des couleurs des pixels d'une image ; définir un algorithme de compression des données et l'organisation de ces données dans un fichier (formats PNG, TIFF...) ; spécifier l'organisation et la structuration d'informations textuelles à partir de l'encodage élémentaire des caractères (formats SGML, XML) ; définir comment les quatre informations élémentaires que sont la numérisation (nombre entier positif), l'espacement (nombre entier positif), le signe de l'exposant et le signe de la mantisse (caractères « + ») sont organisées pour représenter un nombre réel sous forme numérique (cf. standard ANSIMEE 754-1985) (Source : PIAF)
- Pérennité
  - Format de fichier
    - MD
    - Autres contrôles à l'entrée du SAE
    - Intégrité
    - RETEX ETH Zurich
      - Accepter ou pas : comment placer le curseur ?
      - manuellement ?
      - automatiquement ?
      - à l'entrée
      - au cours du cycle de vie ?
      - Transformer ou pas ?
      - contrôle des transformations ?
      - Qui l'assure ?
  - Liabilité
    - Méthodes d'échantillonnage
    - Actions manuelles
    - Comment évaluer la liabilité des fichiers conservés dans le SAE ?
  - Classification des formats
    - Formats stockés
    - Formats identifiés
    - Formats connus
    - Formats maîtrisés
    - Format à l'étude
    - Format potentiellement archivable
    - Format archivable
    - Format en voie d'obsolescence
    - Format obsolète
    - RETEX BnF
    - RETEX CINES
  - Méthodologie de conversion
    - PLATO / PLANETS
    - RETEX CINES : le processus de migration logique est identifié, basé sur la méthodologie PLATO repris dans PLANETS puis OPF :
      - étudier la pertinence de la migration (par format, mais en réunissant les représentants de tous les producteurs) ; nombre de fichiers concernés ; services versants impliqués ; étude du format d'origine ; identification des formats cible ; identification des significatives propriétés à conserver lors de l'opération de migration ; identification des outils. Cette phase est cruciale, avant même la recherche d'outils ;
      - évaluer le temps nécessaire pour mettre en œuvre l'opération de conversion (la conversion des fichiers .pdf archivés dans des versions antérieures à la 1.4., est envisagée depuis des années, mais repoussée, faute de temps) ;
      - réaliser les tests, notamment sur un échantillon (à déterminer sur étude de changement de format) ;
      - évaluer la qualité du résultat ;
      - développer le moyen de réinjecter les nouvelles versions des fichiers dans la plate-forme ;
      - réaliser de manière effective la migration
      - archiver le dispositif et documenter les opérations.
    - L'avantage de la méthodologie mise en œuvre par PLATO est de proposer un arbre de décision basé sur des pondérations.
    - RETEX INA : Notion de filière dans SPAR
      - Trouver de l'homogénéité dans l'hétérogénéité ?
      - les spécifications doivent être accessibles et récupérables pour que le format soit archivable ;
      - l'usage du format doit être largement répandu et le nombre de fichiers à archiver doit être suffisamment important (pas d'investissements sur un format représenté par un très faible nombre de fichiers) ;
      - la conversion est la stratégie de préservation retenue par le CINES, même si l'émulation reste valable pour tout ce qui est animé. Le CINES est conscient que toute migration entraîne des pertes et que, en priorité de toute opération de conversion, il est essentiel de savoir ce qui est le plus important dans le fichier (signifiant propriétés) ;
      - après une opération de conversion, sont conservés l'original, la dernière version migrée et la version migrée, l'identifiant unique du 1er paquet sera toujours conservé ;
    - RETEX CINES
    - RETEX INA
  - Conversion de formats
    - Documentation, mise à jour et pérennisation des référentiels
    - Objectif même du processus de pérennisation ?
    - Impact sur les choix de formats au moment de la collecte ou de l'entrée dans le SAE ?
    - Aucun quel(s) outil(s) ?
    - Pour quels usages ?
    - Accès
    - Audit de conformité OASIS
    - Certification ISO 16363
    - ISO 16919 (pour accréditation)
    - Core Trust Seal
    - Outil d'aide-évaluation en ligne
      - Méthodo DRAMBORA
      - Analyse des risques
      - Curator Costs Exchange
      - Evaluer les coûts
      - 6 - Audit et certification
      - DRAMBORA (2009)
      - NF 42-013
      - TRAC
      - RETEX CINES
      - CoreTrustSeal