



# Le traitement des fonds d'archives

La problématique des  
documents informatiques  
François Cartier

# La problématique

- Apparition de nouveaux supports d'information dans les fonds d'archives acquis
- Obsolescence technologique
  - Appareils pour lire ce qui se trouve sur ces supports
  - Applications pour ouvrir les programmes qui contiennent l'information

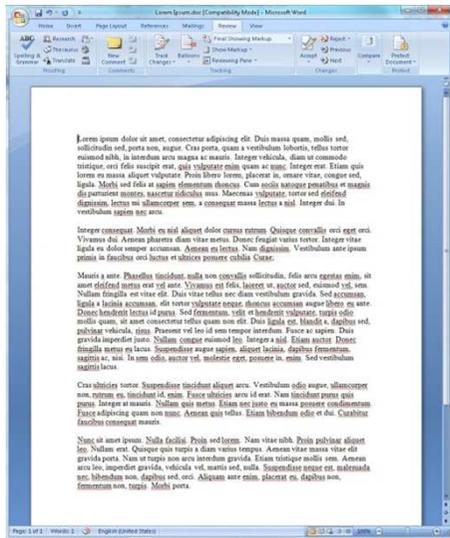
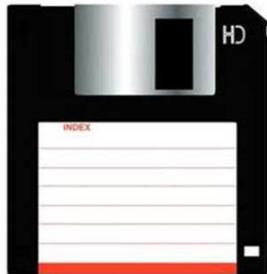
# Le document informatique

- De quoi parle-t-on au juste?
  - **Définition des RDDA** : (...) documents d'archives encodés pour en permettre la manipulation par ordinateur.
  - **Définition de l'EBSI** : Regroupement d'informations ou de données stockées sur support électronique (par exemple, disque optique numérique, ruban magnétique, disquette, etc.).

# Le document informatique

- **Définition de *Techno-science.net*** : *Un fichier informatique est une unité informationnelle physiquement stockée sur un support de mémoire de masse permanent (disquette, disque dur, ruban magnétique...).*
- *Un fichier a généralement comme attributs :*
  - *un nom ;*
  - *un chemin d'accès ;*
  - *une taille mesurée en octets ;*
  - *une date de création et une de dernière modification.*

# Le document informatique



# Dématérialisation



- Retrouver sur un support *informatique* des informations qui :
  - Existent aussi sous forme *analogique* (papier, photo, film, microfilm) = numérisation
  - Sont nées dans un environnement informatique = pas de copie analogique
    - L'information (textuelle, iconographique, sonore...) est représentée sous forme de « bits » (*binary digits*)





# Les principaux supports

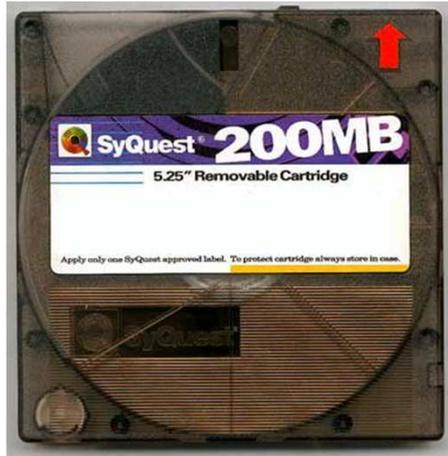
- Seconde génération
  - Les supports **magnétiques** (disquettes, rubans, disques durs); début 1970 à aujourd'hui

8 po.

5 ¼ po.

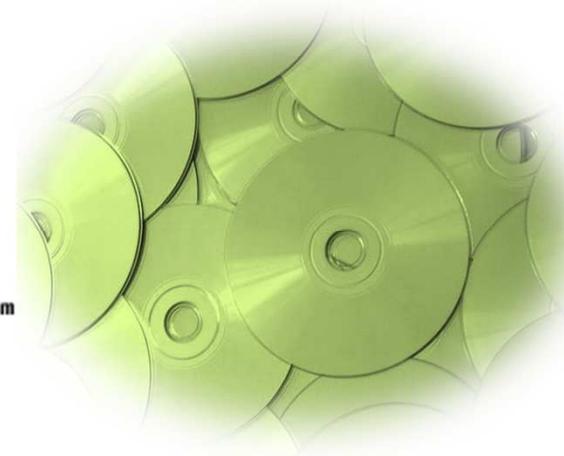
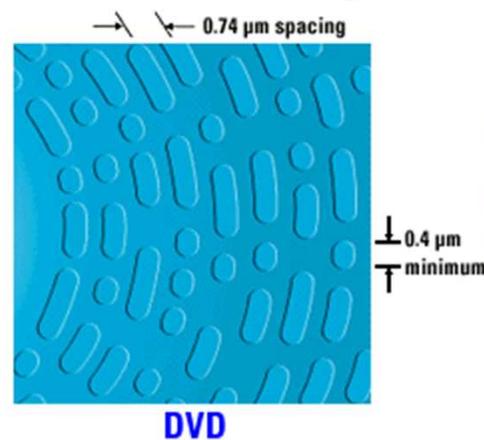
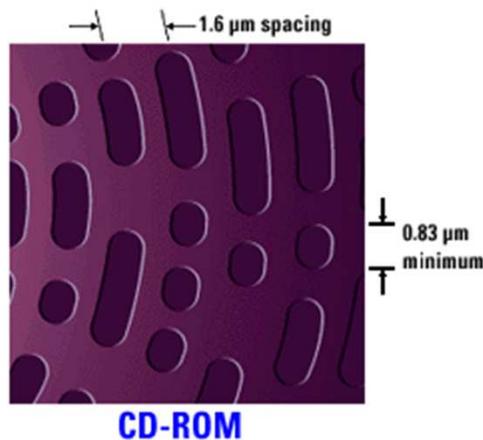
3 ½ po.





# Les principaux supports

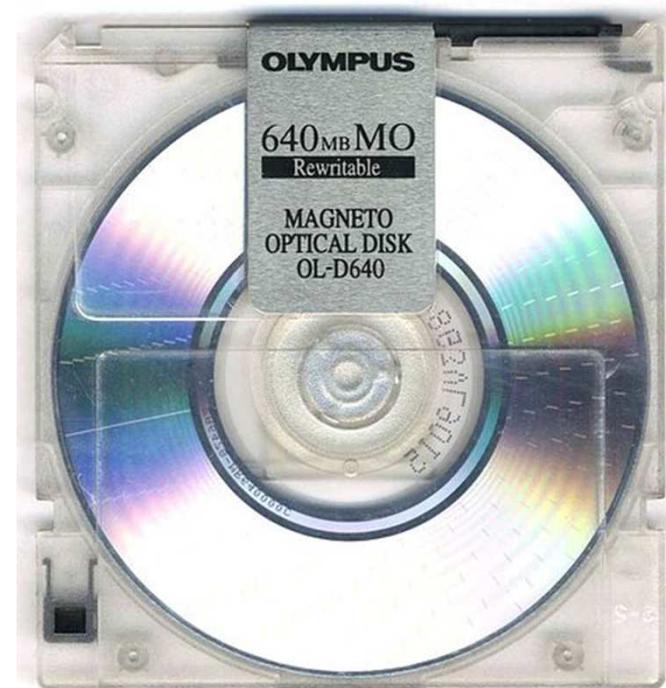
- Troisième génération
  - Les supports **optiques** (cédéroms, DVD, Blu-Ray); 1982 à aujourd'hui



Voir : <http://www.commentcamarche.net/contents/pc/cdrom.php3>

# Les principaux supports

- Troisième génération
  - Les supports **magnéto-optiques**; fin des années 1980 à aujourd'hui
  - Écriture = magnétique
  - Lecture = optique



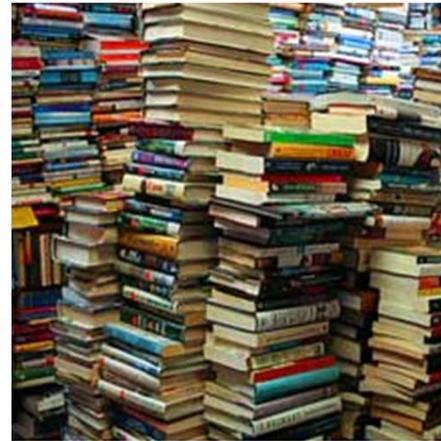
# Les principaux supports

- Quatrième génération
  - Les supports à **semi-conducteurs** ré-inscriptibles (technologie *Flash* : clés USB, cartes SD, etc.); ca. 1985 – aujourd'hui



# Les principaux supports

- Stockage de masse
  - 650 Mo de données = 300,000 pages de texte ( ou 750 livres de 400 pages)



# Les défis de la conservation

Type de supports	Longévité moyenne
<b>Disques magnétiques</b>	
Disques durs	2 à 5 ans
Disquettes	5 à 15 ans
<b>Bandes magnétiques</b>	
Numériques	5 à 10 ans
Analogiques	10 à 30 ans
<b>Disques optiques</b>	
Disques enregistrables	5 à 10 ans
Disques réinscriptibles	5 à 10 ans
Disques enregistrables (couche d'argent métallique)	10 à 50 ans
Disques enregistrables (couche d'or métallique)	100 ans +

# Solutions de conservation

1. Conserver, en même temps que le fichier, **l'ordinateur et les programmes** ayant servi à la production et lecture du document.



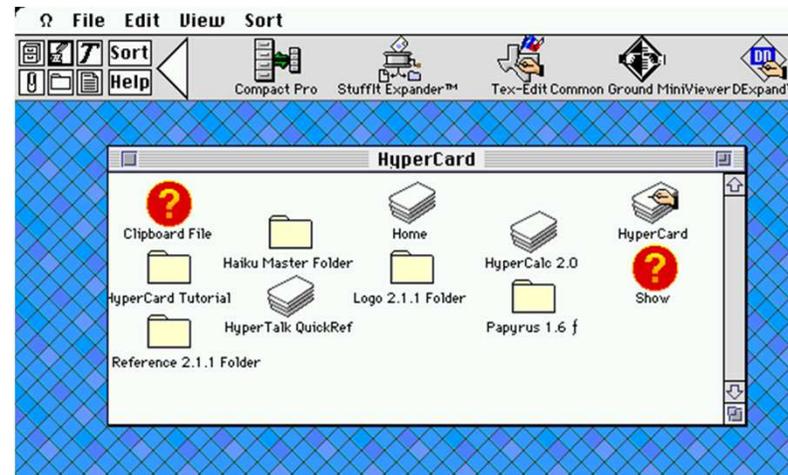
# Solutions de conservation

2. Élaborer des programmes permettant d'**émuler** (imiter) des programmes plus anciens pour lire les fichiers conservés dans leurs formats d'origine.

En informatique, un émulateur est du matériel et / ou un logiciel qui **reproduit** (ou émule) les fonctions d'un système informatique original dans un second système différent, de sorte que le comportement du second système ressemble étroitement au comportement du système original.

# Solutions de conservation

2. Élaborer des programmes permettant d'**émuler** (imiter) des programmes plus anciens pour lire les fichiers conservés dans leurs formats d'origine.



# Solutions de conservation

3. La **migration** des données vers un autre support et un autre format.

## Migration de données

Le transfert périodique des données numériques d'une configuration matérielle et logicielle à une autre ou d'une génération d'ordinateurs à la suivante.

L'objectif de la migration est de **conserver l'intégrité** des documents numériques et de perpétuer la capacité des usagers à les retrouver, les afficher et les utiliser malgré l'obsolescence technologique.



# Solutions de conservation

3. La **migration** des données vers un autre support et un autre format.

Choisir un **support** fiable et durable

- Cédéroms et DVD
- Bandes magnétiques à haute capacité

[www.cci-icc.gc.ca](http://www.cci-icc.gc.ca)



# Solutions de conservation

3. La **migration** des données vers un autre support et un autre format.

Choisir un **format** de données « **ouvert** »

- Images : JPEG2000, PNG, GIF, [TIFF]
- Texte : Office Open XML, Open Document, PDF/A
- Pages web : XML, HTML
- Autres : MNG (film), WAV et AIFF (audio), ZIP (archivage et compression)

# Solutions de conservation

- **Quelques normes ISO liées aux documents informatiques**

- **ISO/TR 10255:2009** Applications de la gestion des documents -- Technologie de stockage sur disque optique, gestion et normes
- **ISO/TR 15801:2009** Images électroniques -- Stockage électronique d'informations -- Recommandations pour les informations de valeur et leur fiabilité
- **ISO 19005-1:2005** Gestion de documents -- Format de fichier des documents électroniques pour une conservation à long terme -- Partie 1: Utilisation du PDF 1.4 (PDF/A-1)
- **ISO 19005-2:2011** Gestion de documents -- Format de fichier des documents électroniques pour une conservation à long terme -- Partie 2: Utilisation de l'ISO 32000-1 (PDF/A-2)
- **ISO/DIS 13008** Information et documentation -- Processus de conversion et migration des documents d'activité numériques (projet)

# Un plan de conservation

- Se doter d'un plan qui :
  - Identifie la nature des documents informatiques dans nos fonds et collections;
  - Identifie leurs spécificités;
    - Format;
    - Système d'exploitation et matériel de lecture;
    - Durée de vie.
  - Met en place des politiques de migration et de tests des supports

# Le traitement des archives

- « Ensemble des opérations et procédures par lesquelles les archives sont triées, classées, décrites et indexées. »  
(Champagne et Chouinard, 1987, p. 163 )



# Le traitement des archives

Étapes du traitement :

1. Collecte
2. Traitement préliminaire
3. Évaluation et tri
4. Classification, classement et cotation
5. Description et indexation
6. Diffusion

# Le traitement des archives

## Traitement préliminaire :

- « [...] opération intellectuelle et matérielle qui assure l'entrée des documents au dépôt. [...]

Cette opération rend tout arrivage communicable aux chercheurs dans un délai raisonnable de quelques jours, de quelques semaines au plus. [...]

(Champagne et Chouinard, 1987, p. 31).

# Le traitement des archives

## Tri :

- Activité archivistique découlant de l'évaluation, qui consiste à appliquer un mode de sélection ou d'échantillonnage.
- Le tri se fait en fonction d'une **politique de tri** institutionnelle qui regroupe des **critères de tri** pour les archives historiques.

# Le traitement des archives

## **Classification :**

- Ensemble de conventions, de méthodes et de règles de procédure structurées logiquement et permettant le classement des documents par groupes ou par catégories, quel que soit leur support et leur âge. »  
(Couture et al., 1999, p. 18).
- Différence entre classification et classement

# Classification

- Classement des archives selon les **activités** du créateur des documents
  - Activités personnelles, relations avec les médias...
- Classement pas **thème / sujet**
  - Pour une collection de photos : photos d'intérieur, de paysages, de bâtiments...
- Classement les archives selon leur **type**
  - Documents textuels, documents cartographiques...

# Classification

- Classement selon la **fonction** des documents
  - Correspondance, documents légaux, cartes, négatifs...
- Classement par ordre **chronologique**
  - Par année, pas décennie, etc.
- Classement **hybride**
  - Reprend deux ou plusieurs des systèmes précédents

# Classification par activités

- Activité « **Voyages d'affaires** » :
  - Billets d'avion et itinéraires sur papier
  - Agenda des rencontres sur format informatique (fichiers Outlook)
  - Photographies numériques des membres des comités rencontrés
  - Cartes (papier) des villes visitées avec annotations



- **Fonds constitué de 12 séries :**

- P6/1 Documents constitutifs;
- P6/2 Direction;
- P6/3 Finances et trésorerie;
- P6/4 Affaires juridiques et légales;
- P6/5 Ressources humaines;
- P6/6 Système de transport en commun de surface;
- P6/7 Système de transport en commun par métro;
- P6/8 Biens immobiliers;
- P6/9 Études;
- P6/10 Imprimés et journaux;
- P6/11 Photographies et autres supports non-textuels;
- **P6/12 Documents informatiques**

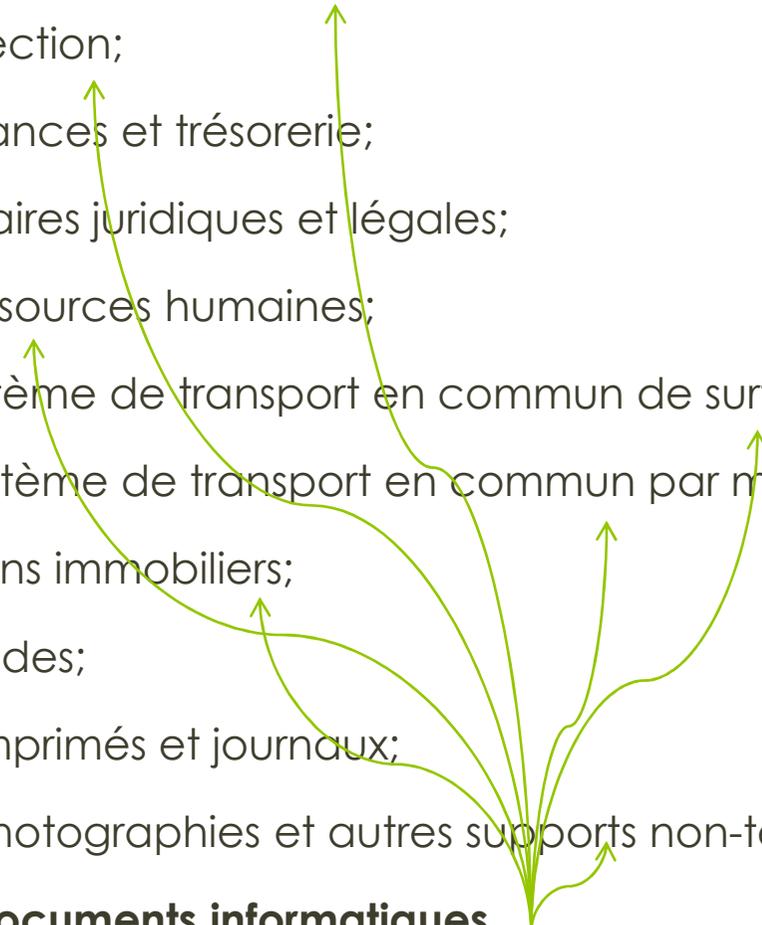


- **Fonds constitué de 12 séries :**

- P6/1 Documents constitutifs;
- P6/2 Direction;
- P6/3 Finances et trésorerie;
- P6/4 Affaires juridiques et légales;
- P6/5 Ressources humaines;
- P6/6 Système de transport en commun de surface;
- P6/7 Système de transport en commun par métro;
- P6/8 Biens immobiliers;
- P6/9 Études;
- P6/10 Imprimés et journaux;
- P6/11 Photographies et autres supports non-textuels;
- **P6/12 Documents informatiques**



- **Fonds constitué de 12 séries :**

- P6/1 Documents constitutifs;
  - P6/2 Direction;
  - P6/3 Finances et trésorerie;
  - P6/4 Affaires juridiques et légales;
  - P6/5 Ressources humaines;
  - P6/6 Système de transport en commun de surface;
  - P6/7 Système de transport en commun par métro;
  - P6/8 Biens immobiliers;
  - P6/9 Études;
  - P6/10 Imprimés et journaux;
  - P6/11 Photographies et autres supports non-textuels;
  - **P6/12 Documents informatiques**
- 
- A diagram consisting of several green arrows originates from the text 'P6/12 Documents informatiques'. These arrows point upwards and to the left, connecting to the following series: P6/10 Imprimés et journaux, P6/9 Études, P6/8 Biens immobiliers, P6/7 Système de transport en commun par métro, P6/6 Système de transport en commun de surface, P6/5 Ressources humaines, P6/4 Affaires juridiques et légales, P6/3 Finances et trésorerie, and P6/2 Direction. The arrow pointing to P6/1 Documents constitutifs is the longest and most direct.



# Le traitement des archives

## **Description :**

- « Enregistrement de l'information portant sur la structure, les fonctions et le contenu des documents présenté sous une forme normalisée. »  
(Bureau canadien des archivistes, 1990, p. D-2).

## **Indexation :**

- « [...] transcrire en langage documentaire les concepts après les avoir extraits du document par une analyse. »  
(AFNOR, 1996a, NF Z 47-102, p. 512).

# Le traitement des archives

- Notices pour un **dossier** comprenant des photos numériques :

P14,S2,D2 – Naissance de bébé. – Novembre 2011. – 3 cm de documents textuels. – 75 photographies (tiff) : coul.. – 1 objet.

ou

P14,S2,D2 – Naissance de bébé. – Novembre 2011. – 3 cm de documents textuels. – 235 Mo de photographies (tiff) : coul.. – 1 objet.

ou

P14,S2,D2 – Naissance de bébé. – Novembre 2011. – 3 cm de documents textuels. – 1 cédérom (75 photographies (tiff) : coul.). – 1 objet.

# Le traitement des archives

- o Notices **à la pièce** pour un document informatique :



P14,S2,D2,P6 – Bébé au repos dans sa couchette. – 14 novembre 2011. – 1 photographie : fichier tiff (4 Mo), coul.; 300ppp.

Portée et contenu : La photo montre un bébé suite à sa naissance à Montréal le 14 novembre 2011. On le voit assoupi, couché sur le dos dans un berceau.

Source du titre propre : Titre basé sur le contenu du document

Variante du titre : IMG0025

Collation : Le fichier est compressé selon l'algorithme LZW. Enregistré sur un cédérom de modèle *Imation 100PK*. Image RGB en 256 couleurs (8 bits).

# Le traitement des archives

- Notices à la pièce pour un document informatique (2) :



P14,S2,D2,P1 – Certificat de naissance. – 14 novembre 2011. –  
1 document textuel : fichier PDF (125 Ko); 24 x 21,5 cm.

Portée et contenu : Le document offre des renseignements...

Variante du titre : CERT-14112011-07

Collation : Le fichier est du type PDF/A-1b et fait deux pages. Il est enregistré sur un cédérom de modèle *Imation 100PK*.

# En conclusion

- Intégration des documents informatiques dans la structure générale du fonds;
- Assurer la pérennité des informations
  - Support stable
  - Langage informatique ouvert
  - Descriptions basées sur les RDDA
- Politiques et procédures



Merci

Questions?