

LE CATALOGUAGE DES MANUSCRITS: DIALOGUE DES FORMATS

De nos jours, les campagnes de numérisation du patrimoine deviennent de plus en plus fréquentes, et différents points sont à considérer comment traiter au mieux un fonds historique concret pour garder son authenticité. L'W D P p O L R U H U V R Q L Q W H U R S p U D E L O L W p " & H U V permet de présenter sous différents formats numériques les métadonnées de sources historiques, mais cette diversité incite toujours à créer des modèles multiples pour leur traitement. La question G ¶ X Q H Q R U P D O L V D W L R Q G X, afin de pouvoir communiquer entre eux, différentes institutions scientifiques de manière globale et de pousser donc à une meilleure interopérabilité des données.

À l'exemple des fonds anciens de la Bibliothèque de Saint-Omer, contenant O H V P D Q X V F U L W V P p G L p Y B R X, cet article traite de la sélection de solutions proposées pour la conversion des formats. La problématique principale concerne la conversion des métadonnées dans différents formats en gardant leurs spécificités et en faisant D W W H Q W L R Q j O ¶ H Q A F E R S D U R A T I O N L e s u j e t p r é s e n t é s e r a d é v e l o p p é s en trois étapes: la contextualisation du projet Saint-Bertin; O ¶ p Y R F D e v a i n s Q u e s t i o n s sur les formats numériques; les extraits du schéma de conversion illustrant les différents objectifs des formats, ce qui a finalement constitué un défi numérique pour la conversion du catalogue de Saint-Bertin.

Le projet «Saint-Bertin, centre culturel du VI^e au XVIII^e siècle: constitution, conservation, diffusion, utilisation du savoir H V W P H Q p G H S X L V r H F S K D U F o n d s d e l a B i b l i o t h è q u e d e s t e x t e s (I R H T C N R S), O ¶ e T X L S H [% L E D O L R W a g g e m e n t d e l a B i b l i o t h è q u e d e S a i n t - O m e r (B A S O) e t l a l i b r a i r i e d e B o u l o g n e - M e r (B S M) . C e p r o j e t f a i t p a r t i e d e m a i n t s p r o j e t s numériques français de nos jours dont le but est la reconstitution des bibliothèques médiévales

¹ Cette conversion a été réalisée au code X Q D J H D X V H L Q G H O D V H F W L R Q, G K L O D ¶ B D W p R. inscrit dans le projet Saint-Bertin, centre culturel du VI^e

fusion de toutes les métadonnées et leur enrichissement, normalisation, standardisation et

FUpDWLRQ GHV OLHQV GDQV OHV EXWV G¶LQWHURSpUDELO

Tableau1: Structure de la partie concernant le manuscrit (<msDesc>)

La structure choisie semble être précisée et adaptée à la création du catalogue électronique des
PDQXVFULWV GHV IRQGV DQFLHQV /H UHJURXSHPHQW GHV
<teiHeader> (dans la structure du fichier, elle est un élément nécessaire et se trouve tout au début)
et toutes les notices de manuscrits se trouvent dans <text> qui représente une liste contenant
plusieurs <msDesc> FH TXL DVVXUH O¶HUJRQRPLH O¶H[SORLWDWLF
fichier.

Certes, la structure globale du fichier TEI (avec la division en

	<pre> </dsc> </archdesc> </ead> </pre>
--	--

Tableau 2: \$ U E R U H V F H Q F (gauche) / Q ¶ \$ \$ ' H Q F R G (droite) /

On voit sur le Tableau 2 à gauche que OD VWUXFWXUH GH O ¶ (\$ De p m s W U q V sur le même tableau à droite elle est comparable avec celle de la TEI Si, dans la TEI la division globale se présente comme `<teiHeader>` (métadonnées sur le document électronique) et `<eadHeader>` (métadonnées sur le contenu).

GRQQ p HV VXU OHV REMHWV HQFRG p V <eadHeader> (pour les F H W W H métadonnées) <archdesc> (pour la description du contenu). / ¶ p O p F i d O W i t t e r > ¹⁶ Q ¶ H V W S D V R E O L J D W R L U H P D L V L O S H U P H W G ¶ H Q F R G H U O H V L Q I R U P D W G p F U L W V T X L V H W U R X Y H Q W G D Q V O H ¹⁷ C o m p o s a Q s W U H Q F R G p f o n d s) q u i p e u v e n t d e s H Q G U H G D Q V O D u d e t i n e L D s i e f i l i e r s l e q u e l s v o t r e f i e n t s o n t c o n s e r v é s .

Néanmoins, la Bibliothèque nationale de France utilise le format EAD également pour O ¶ H Q F R G D J H G H V C a t a l o g u e g é n é r a l d e s m a n u s c r i t s l i b r a i r e s p u b l i c s d e F r a n c e (CGMBPF ou plus raccourci CGM) existe depuis 1838 et est passé depuis 1993

¹⁵ Récupérée de Guide pour la mise à jour des fichiers issus de la conversion rétrospective du CGM http://www.bnf.fr/documents/cgm_retroconverti_mise_a_jour.pdf

¹⁶ &HW p O p P H Q W G R Q Q H G H V L Q I R U P D W L R Q V V X U O D F U p D W L R Q O D S X E lui-même et non sur les documents en cours de description. Encoded Archival Description, Tag Library, Description Archivistique Encodée, Dictionnaire des balises R F L H W \ R I \$ P H U L F D Q \$ U F K L Y L V W V W le groupe AFNOR CG46/C N357/GE3, octobre 2004, p. 137, <http://www.archivesdefrance.culture.gouv.fr/static/1066>

niveau pa

préférerenirles donnéesG ¶ X Q H P D à partir de la notice de requête. Cette divergence témoignedes G L I I p U H Q W V R E M H F W L I V G H F H V G H X [I R U P D W V G D Q V

La multiplicité des options de conversion permises par le nombre des balises existantes dans O D 7 (, H W S O ¶ X ¶ Q Y D V W H F K D P S G ¶ K \ S R W ¶ s partitularités des S H U p I O H formats epousse les institutions présentant le produit définitif à limiter leur choix de format utilisateurs et des enjeux X X ¶ L O V U H S U p a i l l e d e c o n v e r s i o n aide également à pointer le format à partir duquel on est H F W X H F H W W H F R Q Y H U V L R i e d o d r e c o n v e r s i o n être bien structuré et encodé de façon logique qui se répercutera sur une extraction rapide des données

De même le traitement de chaque information visant à conserver celles qui sont importantes reste une des principales questions Comment mener au mieux le dialogue entre le créateur, I ¶ X W L O L V D W H X U C e r t a i n e s s o l u t i o n s d e d o c u m e n t s p r é s e n t é e s i c i v i s e n t à respecter la structure du document de sortie en EAD à garder la logique de la notice très enrichie en TEI.

Annexes

1. Métadonnées

<p><teiHeader></p>	<p><eadheader></p>
<p></teiHeader></p>	<p></eadheader></p>

2. Changement de structure

--	--

5. *Histoire du manuscrit*

<pre><history> <origin> </origin> <provenance></pre>	<pre>unitdate </unitdate> <custodhist> TEXTE </custodhist></pre>
--	--