

Bibliothèques et documents numériques

*Concepts, composantes,
techniques et enjeux*

par Alain Jacquesson
et Alexis Rivier

Collection
Bibliothèques

ÉDITIONS
DU CERCLE DE LA LIBRAIRIE

Table

Avant-propos	11
Introduction	13
PARTIE I - LES PARADIGMES	21
1. La chaîne contemporaine de production des documents	23
1.1 Les logiciels de création des documents	23
1.2 De l'auteur à l'éditeur sans rupture	24
2. La représentation numérique des documents	29
2.1 La représentation numérique des textes	30
2.1.1 Le code Ascii	32
2.1.2 Ascii et la représentation multiscrite	33
2.1.3 Unicode	34
2.2 La représentation numérique des images	36
2.3 Les techniques de compression des informations	40
3. L'architecture des documents	43
3.1 La structure des documents	44
3.1.1 Structure logique et structure physique	44
3.1.2 Structure générique et structure individuelle	47
3.2 Les langages de balisage des documents	48
3.2.1 Tex	48
3.2.2 ODA — <i>Open Document Architecture</i>	49
3.2.3 SGML — <i>Standard Generalized Markup Language</i>	50
3.2.3.1 Le succès de SGML	50
3.2.3.2 DTD — Définition de type de document	51
3.2.3.3 Les logiciels d'édition	53
3.2.3.4 Les logiciels de visualisation et de navigation	54
3.2.3.5 La présentation des documents SGML	55
3.3 Les liens hypertextes et hypermédias	55
3.4 Le balisage des documents multimédias : HTML	60
3.5 XML — <i>Extensible Markup Language</i>	61
3.6 L'authentification des documents numériques	62
3.7 Les livres électroniques	63
3.7.1 Les fonctionnalités du livre électronique	64
3.7.2 Le livre électronique portable	65

4. Les techniques actuelles de numérisation	71
4.1 Les techniques d'acquisition	71
4.1.1 La saisie au clavier	73
4.1.2 Les scanners bureautiques	73
4.1.3 Les scanners de livres	76
4.1.4 Les appareils photographiques numériques	80
4.1.5 Les scanners de microformes	81
4.1.6 Les scanners de diapositives et de transparents	82
4.1.7 Le passage du mode image en mode texte (OCR/ICR)	85
4.2 Les techniques de conservation	88
4.2.1 Les besoins de stockage	89
4.2.2 L'exploitation des données	91
4.3 Les techniques de restitution	94
4.3.1 Les écrans de visualisation	94
4.3.2 Les imprimantes	96
 PARTIE II – LA CRÉATION DES DOCUMENTS NUMÉRIQUES	 101
5. La production de journaux électroniques	103
5.1 Les supports des périodiques électroniques	105
5.2 Les limites du support papier	106
5.3 La typologie des périodiques électroniques	106
5.3.1 Les revues scientifiques	107
5.3.2 Les revues de sommaires	108
5.3.3 Les lettres d'information	109
5.3.4 Les listes de discussion	110
5.3.5 Les revues « mouvantes »	110
5.3.6 Les Web'zines	111
5.3.7 Les autres types de périodiques électroniques	112
5.4 Les périodiques électroniques contrôlés	112
5.5 Les publics des périodiques électroniques	113
5.6 L'économie des périodiques électroniques	114
5.6.1 Les périodiques électroniques gratuits	114
5.6.2 Les périodiques électroniques payants	114
5.6.2.1 Les éditeurs commerciaux	114
5.6.2.2 Les organisations sans buts lucratif	115
5.6.2.3 Les principes de facturation des périodiques électroniques	115
5.6.3 La publicité	116
5.6.4 Les périodiques électroniques financés par les auteurs	117
5.6.5 Une maison d'édition créée par les bibliothécaires : HighWire Press	117
5.7 La naissance des périodiques électroniques	120
5.7.1 Le projet EIES (<i>Electronic Information Exchange System</i>)	121
5.7.2 Le projet Blend	123
5.7.3 Le projet Adonis	123
5.7.4 Le projet Quartet	124
5.8 Les réalisations commerciales	125
5.8.1 Elsevier	125
5.8.1.1 Le projet Tulip	126
5.8.1.2 Le projet de l'université de Tilburg (Pays-Bas)	130
5.8.1.3 Les abonnements aux périodiques électroniques d'Elsevier	132

5.8.1.4 Les autres produits électroniques d'Elsevier	133
5.8.1.5 La position dominante d'Elsevier	134
5.8.2 Academic Press	135
5.8.3 Springer Verlag et le service Link	136
5.8.4 H. W. Wilson	138
5.8.5 Blackwell's	138
5.8.6 John Wiley & Sons	139
5.8.7 Masson	140
5.8.8 Les Éditions de physique	140
5.9 Les réalisations des universités et des centres de recherche	141
5.9.1 Mathématiques	142
5.9.2 Droit	144
5.9.3 Les revues électroniques en sciences humaines	144
5.9.4 Les journaux d'information universitaire	145
5.9.5 Les journaux d'étudiants	145
5.10 Les sociétés savantes	146
5.10.1 Institute of Physics — IoP	146
5.10.2 Association for Computing Machinery — ACM	148
5.11 Les périodiques électroniques et les bases de données brutes	153
5.12 Les valeurs liées à la fonction d'archivage	156
5.13 L'architecture générale des serveurs de périodiques électroniques	157
6. Les réservoirs de textes numériques	163
6.1 La valeur ajoutée des textes numériques	164
6.2 La normalisation des textes numériques	166
6.2.1 La <i>Text Encoding Initiative</i> — TEI	167
6.2.1.1 <i>Corpus Encoding Standard</i> — CES	170
6.3. Les grands projets de numérisation de textes	171
6.3.1 Les réalisations anglo-saxonnes	171
6.3.1.1 Le projet Gutenberg aux États-Unis	171
6.3.1.2 L'Electronic Text Center (ETC) de l'Université de Virginie	173
6.3.1.3 La bibliothèque universelle de l'Université Carnegie Mellon	175
6.3.1.4 Le Center for Electronic Texts in the Humanities — CETH	176
6.3.1.5 The Oxford Text Archive — OTA	176
6.3.1.6 L'ARTFL de Chicago	176
6.3.1.7 Le Center for Electronic Text in the Law de Cincinnati — CETL ..	178
6.3.1.8 Les autres réalisations anglophones	178
6.3.2 Les réalisations francophones	179
6.3.2.1 L'Association des bibliophiles universels — ABU	179
6.3.2.2 La base Frantext	180
6.3.2.3 Le <i>Trésor de la langue française</i>	181
6.3.2.4 Le projet Athena à Genève	181
6.3.2.5 La bibliothèque électronique de Lisieux	182
6.3.3 Les autres réalisations	183
6.4 Les réservoirs de textes numériques : premier bilan	184
6.5 Les aspects bibliothéconomiques de la gestion des textes numériques	185
6.5.1 Les catalogues collectifs	185
6.5.2 Un nouveau volet du métier	186

7. La numérisation des anciennes collections	189
7.1 Les objectifs des projets de numérisation	189
7.2 La gestion d'un projet de numérisation	191
7.3 La numérisation des microformes	196
7.4 Les grandes réalisations	198
7.4.1 États-Unis	198
7.4.2 Australie	200
7.4.3 France	201
7.5 La coordination des politiques de numérisation	202
 PARTIE III – LES BIBLIOTHÈQUES FACE AUX DOCUMENTS NUMÉRIQUES	205
8. L'organisation des bibliothèques numériques	207
8.1 Les préalables pédagogiques et psychologiques	208
8.2 L'environnement architectural de consultation des collections numériques ..	209
8.3 Les considérants économiques	210
8.4 L'environnement informatique	211
8.5 L'entreposage des collections numériques	211
8.5.1 Les collections numériques matérielles	212
8.5.1.1 Les CD-Roms en prêt	212
8.5.1.2 Les CD-Roms en ligne	212
8.5.2 Les collections numériques en ligne	215
8.5.2.1 Les collections chargées localement	215
8.5.2.2 Les collections accessibles à distance	216
8.6 Repenser l'organisation des bibliothèques	217
 9. Le traitement bibliographique des documents numériques	219
9.1 Le catalogage et la classification des documents numériques	221
9.2 Les méthodes de localisation (URL, URN, PURL, Handle)	225
9.3 Un produit documentaire à valeur ajoutée : les métadonnées	229
9.4 La notion de notice SuperMarc	232
9.5 Les relations entre les modules d'accès publics et les documents numériques	234
 10. Les mécanismes de repérage des documents numériques	239
10.1 Les moteurs de recherche	240
10.1.1 AltaVista	242
10.1.2 Les autres moteurs de recherche	243
10.2 Les répertoires	244
10.2.1 Yahoo!	245
10.2.2 Les autres répertoires	247
10.3 Les méta-index	248
10.4 Comment être référencé dans les moteurs de recherche et les répertoires ...	249
10.5 Les listes établies par les bibliothèques	252
10.6 Les logiciels d'accès public des bibliothèques	254
10.7 L'information personnalisée et les techniques « push »	258
10.8 Citer des documents électroniques	259
 11. Les nouvelles méthodes de lecture	265
11.1 La gestion du document	266
11.2 Les recherches de la Bibliothèque nationale de France	266

11.3 Les fonctionnalités de lecture	268
11.3.1 La recherche et la sélection	268
11.3.2 L'appropriation des textes et la constitution de corpus personnels	269
11.3.3 Les annotations	269
11.3.4 L'exploration linguistique	272
11.3.5 Le classement et l'archivage individualisés	275
12. Les collections spéciales	279
12.1 La description des documents numérisés spéciaux	280
12.1.1 Les collections spéciales et les métadonnées	280
12.1.2 L'intégration des collections spéciales dans le catalogue	282
12.2 Les problèmes particuliers de numérisation	287
12.3 Les collections iconographiques	289
12.3.1 Traitement bibliothéconomique ou muséographique ?	291
12.3.2 Les réalisations	295
12.4 Les manuscrits et les archives	297
12.4.1 La numérisation en mode texte des inventaires – <i>L'Encoding Archival Description (EAD)</i>	299
12.4.2 La numérisation en mode image des documents primaires	302
12.5 Les papyrus	303
12.6 Les affiches	304
12.7 Les cartes et plans	304
12.8 Les collections audiovisuelles	308
13. La préservation des collections numériques	313
13.1 La problématique vue par la <i>Commission on Preservation and Access</i>	313
13.1.1 Les caractéristiques de l'information numérique	314
13.1.2 Les stratégies de conservation de l'information numérique	317
13.1.3 L'organisation des archives numériques	318
13.2 La préservation des périodiques numérisés — JSTOR	318
13.3 La préservation des documents d'origine numérique — Le dépôt légal	320
13.3.1 Dépôt légal et documents électroniques matériels	321
13.3.2 Le dépôt légal et les documents électroniques en ligne	322
13.4 La responsabilité des bibliothèques	326
14. Les implications intellectuelles des bibliothèques numériques	329
14.1 Une réponse aux besoins en éducation ?	330
14.2 La valeur scientifique de la publication électronique	331
14.3 L'intégrité de la publication scientifique	334
14.4 La censure et la surveillance des publications électroniques	335
14.5 Le changement du système de communication scientifique	337
14.6 Les collègues invisibles et la littérature grise	340
14.7 Le droit d'auteur et l'accès aux documents	342
14.7.1 Les principes généraux du droit d'auteur	342
14.7.2 Les licences contractuelles	344
14.7.3 Le point de vue des fournisseurs d'information	346
14.7.4 Le point de vue des bibliothèques	347
14.7.5 Les travaux législatifs en cours	349

15. Mythes et limites des bibliothèques numériques	355
Liste des sigles et des acronymes	361
Index	369