

TABLE DES MATIÈRES

SOMMAIRE	5
PRÉFACE	7
INTRODUCTION	9
Objectifs	9
Structure du livre	11
Limites	13
CHAPITRE 1	
ACQUÉRIR L'INFORMATION	15
Sources d'information et question de recherche	16
1.1.1 <i>Typologie des sources d'information</i>	16
1.1.2 <i>Définition d'une question de recherche</i>	18
Requêtes en langage documentaire	19
1.2.1 <i>Qu'est-ce qu'un langage documentaire ?</i>	19
1.2.2 <i>Définir et articuler les concepts</i>	19
1.2.3 <i>Évaluer les résultats</i>	20
Internet	21
1.3.1 <i>Internet comme infrastructure de base</i>	22
1.3.2 <i>Vecteur d'innovation</i>	23
1.3.3 <i>Architecture client-serveur</i>	24
1.3.4 <i>Importance des protocoles</i>	26
1.3.5 <i>Hypertexte — nouveau paradigme d'écriture et de lecture</i>	27

UX
S :
es
ciales

land
Gillet
chen
Vilde

inatra

NOTO
VERSION NUMÉRIQUE

1.4	Le Web	28
1.4.1	HTTP : transfert de contenus	29
1.4.2	URLs : où aller sur le Web ?	29
1.4.3	HTML : un langage pour le Web	31
1.4.4	Navigateurs web : interpréter le HTML	34
1.5	Moteurs de recherche	35
1.5.1	Crawling et ranking	36
1.5.2	Utilisation pratique des opérateurs de recherche dans Google	37
	En résumé	38
	Pour aller plus loin	39
	Exercices	40

CHAPITRE 2

	MODÉLISER LES DONNÉES	41
2.1	Découvrir différentes visions du monde	42
2.2	Données tabulaires	44
2.2.1	Modèle	45
2.2.2	Implémentation	45
2.2.3	Recherche	46
2.2.4	Évolution	47
2.2.5	Partage	47
2.3	Bases de données relationnelles	48
2.3.1	Modèle	48
2.3.2	Implémentation	51
2.3.3	Recherche	53
2.3.4	Changement	53
2.3.5	Partage	54
2.4	XML	55
2.4.1	Modèle	56
2.4.2	Implémentation	56
2.4.3	XML Schema	60
2.4.4	Namespaces	60
2.4.5	Recherche	62
2.4.6	Perspectives divergentes sur le XML	62
2.4.7	Évolution	63
2.4.8	Partager	64
2.5	Web des données	65
2.5.1	Modèle : RDF	65

2.5.2	Décrire le monde : l'application	
2.5.3	Le pragmatisme des Link	
2.5.4	Sérialisation	
2.5.5	Recherche	
2.5.6	Évolution	
2.5.7	Partage	
2.6	Continuum entre la recherche	
2.6.1	Google Scholar	
2.6.2	Scopus	
	En résumé	
	Pour aller plus loin	
	Exercices	

CHAPITRE 3

	NUMÉRISER LES SOURCES	
3.1	Le patrimoine mondial accessible	
3.1.1	Mythe de l'accès numérique	
3.1.2	Gestion du patrimoine : la	
3.1.3	Vision complète des étapes	
3.2	Préparation	
3.2.1	Que veulent les utilisateurs	
3.2.2	Financement	
3.2.3	Numériser en interne ou	
3.3	Processus de numérisation	
3.3.1	Bases techniques	
3.3.2	Outils de gestion	
3.3.3	Pourtant, des problèmes	
3.4	Métadonnées	
3.4.1	Que sont les métadonnées	
3.4.2	Contexte et évolution	
3.4.3	Standards de métadonnées	
3.4.4	Dénominateur commun :	
3.4.5	Notices d'autorité et vocab	
3.4.6	Méthode de construction	
3.4.7	Problèmes liés aux vocab	
3.4.8	Création automatisée de	
3.4.9	Le crowdsourcing	

2.5.2	<i>Décrire le monde : l'approche des ontologies</i>	67
2.5.3	<i>Le pragmatisme des Linked Data</i>	69
2.5.4	<i>Sérialisation</i>	70
2.5.5	<i>Recherche</i>	71
2.5.6	<i>Évolution</i>	72
2.5.7	<i>Partage</i>	73
	Continuum entre la recherche non structurée et structurée	73
2.6.1	<i>Google Scholar</i>	74
2.6.2	<i>Scopus</i>	75
	En résumé	77
	Pour aller plus loin	78
	Exercices	78

CHAPITRE 3

	NUMÉRISER LES SOURCES	81
	Le patrimoine mondial accessible en un clic ?	83
3.1.1	<i>Mythe de l'accès numérique au patrimoine culturel</i>	84
3.1.2	<i>Gestion du patrimoine : tensions entre passé et présent</i>	85
3.1.3	<i>Vision complète des étapes de la numérisation</i>	87
	Préparation	87
3.2.1	<i>Que veulent les utilisateurs ?</i>	88
3.2.2	<i>Financement</i>	89
3.2.3	<i>Numériser en interne ou en externe ?</i>	91
	Processus de numérisation	93
3.3.1	<i>Bases techniques</i>	93
3.3.2	<i>Outils de gestion</i>	96
3.3.3	<i>Pourtant, des problèmes surviennent</i>	98
	Métadonnées	99
3.4.1	<i>Que sont les métadonnées ?</i>	99
3.4.2	<i>Contexte et évolution</i>	100
3.4.3	<i>Standards de métadonnées et voie vers l'interopérabilité</i>	102
3.4.4	<i>Dénominateur commun : le Dublin Core</i>	105
3.4.5	<i>Notices d'autorité et vocabulaires contrôlés</i>	106
3.4.6	<i>Méthode de construction d'un thésaurus</i>	107
3.4.7	<i>Problèmes liés aux vocabulaires contrôlés</i>	109
3.4.8	<i>Création automatisée de métadonnées</i>	111
3.4.9	<i>Le crowdsourcing</i>	113

UX
S :
es
iales

land
Gillet
chen
Vilde
inatra

NOTO
VERSION NUMÉRIQUE

3.5 Outils de gestion de collection	114
3.5.1 Relations houleuses avec les outils	115
3.5.2 Comment être prisonnier de son logiciel : un exemple du terrain	116
3.5.3 Évaluer la qualité d'une base de données documentaire	117
3.5.4 Omeka, un exemple d'application pour décrire des contenus	120
En résumé	121
Pour aller plus loin	122
Exercices	123

CHAPITRE 4

ANALYSER LE CONTENU	125
4.1 Comprendre un contenu textuel	127
4.1.1 Traitement automatique des langues	127
4.1.2 L'ère des Big Data	130
4.2 Lire un texte à distance	131
4.2.1 Un concordancier : AntConc	131
4.2.2 Analyse lexicale avec Unitex	134
4.3 Constituer son corpus de recherche	136
4.3.1 Scraping	137
4.3.2 Étude de cas : les discours de la présidence Hollande	138
4.3.3 Nettoyer ses données avec OpenRefine	140
4.4 Enrichir ses données	142
4.4.1 Extraction d'information	142
4.4.2 Topic modeling	145
4.5 Faire parler ses données : techniques de visualisation	146
En résumé	147
Pour aller plus loin	148
Exercices	149

CHAPITRE 5

VALORISER LES RÉSULTATS	151
5.1 Outils de rédaction scientifique	152
5.1.1 LaTeX	153
5.1.2 Markdown	154
5.1.3 Rédiger en collaboration	154
5.1.4 Gestion des références bibliographiques	155
5.2 Création d'un carnet de recherche en ligne	158
5.2.1 Encoder en HTML	159

5.2.2	Fonctionnement des CMS.....	161
5.2.3	Mettre ses contenus en ligne grâce à Omeka.....	164
3	Hébergement : terre à terre ou dans le cloud ?.....	167
5.3.1	Contexte historique.....	168
5.3.2	Principes et impact du cloud computing.....	170
4	Assurer la visibilité des contenus.....	171
5.4.1	Spamming.....	173
5.4.2	Achat d'annonces publicitaires.....	175
5.4.3	Calcul de l'impact des contenus en ligne.....	177
5.4.4	Problématique des cookies.....	180
5	Valoriser dans le respect des droits.....	182
5.5.1	Le droit d'auteur.....	183
5.5.2	Le droit à l'image.....	185
5.5.3	Protection de la vie privée.....	185
5.5.4	Open source : entre utopie et commerce.....	186
	En résumé.....	187
	Pour aller plus loin.....	188
	Exercices.....	189
	CONCLUSIONS.....	191
	BIBLIOGRAPHIE.....	195
	INDEX.....	199

UX
S :
es
iales
land
Gillet
chen
Wilde
inatra