

**Solutions
développeurs**

D. HUNTER,
C. CAGLE, D. GIBBONS, N. OZU, J. PINNOCK, P. SPENCE

Initiation à XML

Avec trois études de cas détaillées



Au sommaire

- Syntaxe XML et écriture de documents bien formés
- Feuilles de style CSS et transformations XSLT
- Navigation et liens avec XPath, XPointer et XLink
- Traitement de documents XML avec les API DOM et SAX
- Validation avec les DTD et les schémas XML
- XML et les bases de données
- Études de cas : site marchand, groupe de discussion, application B-to-B

www

Sur le site www.wrox.fr

- Téléchargez le code des exemples
- Abonnez-vous à notre liste [coup2main.wrox.fr](mailto:coup2main@wrox.fr)
- Bénéficiez du support technique gratuit

E Eyrolles

wrox

2-005-281

Initiation à XML

Avec trois études de cas détaillées

David Hunter

avec la collaboration de

Curt Cagle • Dave Gibbons • Nikola Ozu • Jon Pinnock • Paul Spencer

Traduit de l'anglais par Fabrice Lemaingue, Laurent Adam, Clotilde Raspaud



Eyrolles



Table des matières

Introduction

À qui s'adresse ce manuel ?	1
De quoi traite ce manuel ?	2
De quoi avez-vous besoin pour ce manuel ?	2
Conventions utilisées	5
Télécharger le code source	6
Votre avis nous intéresse	7
Errata & mises à jour	7

Chapitre 1 : Qu'est-ce-que le XML ?

À propos des données, des fichiers et du texte	9
Fichiers binaires	10
Fichiers texte	10
Historique des langages de balisage	11
XML, qu'est-ce que c'est ?	12
Qu'est-ce que XML a de plus ?	13
Comment décrire autrement les données ?	16
Parseurs XML	17
Pourquoi « Extensible » ?	18
HTML et XML : le gâteau et la cerise sur le gâteau	19
Hiérarchie de l'information	20
Hiérarchies et HTML	20
Hiérarchies et XML	20
Qu'est-ce qu'un type de document ?	22
Sérieusement... Qu'est-ce qu'un type de document ?	23

Qu'est-ce que le World Wide Web Consortium ?	24
Quels sont les composants de XML ?	25
Utilisation réelle et potentielle de XML	26
Réduction de la charge de travail du serveur	26
Contenu de site web	26
Appel de procédure à distance (RPC)	26
Commerce électronique	26
Résumé	27

Chapitre 2 : Écrire du XML bien formé 29

Balises, texte et éléments	30
Règles applicables aux éléments	34
Toute balise d'ouverture doit posséder une balise de fermeture correspondante	34
Les balises ne peuvent pas se chevaucher	35
Un document XML ne peut posséder qu'un seul élément racine	36
Noms d'élément	36
Sensibilité à la casse	37
Espaces vides dans des PCDATA	38
Attributs	41
Pourquoi utiliser des attributs ?	45
Commentaires	46
Éléments vides	50
Déclaration XML	51
Encodage (attribut encoding)	52
Unicode	52
Spécifier l'encodage de caractères pour XML	53
Standalone	53
Instructions de traitement	55
La déclaration XML est-elle une instruction de traitement ?	56
Caractères PCDATA illégaux	58
Utiliser des caractères d'échappement	58
Sections CDATA	59
Analyser XML	62
Parseur Internet Explorer (Microsoft)	62
Expat (James Clark)	63
ActiveDOM (Vivid Creations)	63
DataChannel XJ Parser	63
xml4j (IBM)	63
Xerces (Apache)	63
Libxml / gnome-xml	63
Erreurs en XML	63
Résumé	64

Chapitre 3 : XML dans le navigateur : feuilles de style en cascade

	65
Pourquoi les feuilles de style sont-elles nécessaires ?	66
Historique des styles HTML	67
Origines du balisage	68
Feuilles de style et HTML	70
Créer des styles en HTML	72
L'attribut style	72
L'élément <STYLE>	72
Styles en cascade	73
Écraser les styles avec des éléments STYLE	76
Feuilles de style et commentaires	77
Utiliser des classes en HTML	77
Sélecteurs de classe en cascade	81
Pseudo-classes	84
Styles externes en HTML	87
<LINK>	87
Directive @import	88
Feuilles de style en cascade et XML	89
Feuilles de style externes en XML	89
Plus de détails sur les feuilles de style externes	92
Feuilles de style en cascade XML internes	93
Le média est le message	94
Directive @media	94
Groupes de médias	95
Le modèle de boîtes	96
Propriété display	97
Autres valeurs de display	98
Différence entre display et visibility	99
Définir des couleurs	101
Codes de couleur hexadécimaux	102
Dessiner des bordures	103
Images d'arrière-plan	107
Limites des manipulations d'images	108
Positionner et répéter des images d'arrière-plan	108
Spécifier plusieurs propriétés d'arrière-plan	109
Créer un arrière-plan	110
Comprendre les schémas de positionnement	112
Position static	112
Position relative	112
Position absolute	112
Position fixed	113
Héritage de position	113
Définir les positions et les unités	113
Propriétés width et height (largeur et hauteur)	115
Propriété overflow (débordement)	116

Propriété clip (découpe de la visualisation)	116
Propriété float (éléments flottants)	117
Listes et tableaux	119
Travailler avec des listes en XML	121
Prendre en charge les tableaux en CSS	123
Manipulation de texte et de polices	125
Propriété font-size (taille de police)	125
Propriétés font-style et font-weight (style et graisse de police)	126
Propriété font-family (familles de polices) et directive @font-face	127
Propriétés font-variant, text-decoration, text-transform	128
Indentation, espacement et marges	128
text-indent	128
Marges et espacement	129
Crénage et espacement entre caractères	130
letter-spacing	130
line-height	130
Alignement	131
text-align	131
vertical-align	131
Caractères Ruby	131
Impression	132
Unités : pixels contre points	132
Directive @page	133
Saut de page	134
Impression à l'aide de scripts	134
Feuilles de style aurales	134
Résumé	137

Chapitre 4 : XSLT et XPATH

	139
Exécuter les exemples	140
XT	140
MSXML	140
Qu'est-ce que XSL ?	141
Comment fonctionnent les feuilles de style XSLT ?	142
Associer des feuilles de style à l'aide d'instructions de traitement	143
Pourquoi XSLT est-il si important pour le commerce électronique ?	143
Programmation impérative et programmation déclarative	146
Programmation impérative	146
Programmation déclarative	147
XSLT n'a pas d'effet de bord	149
XPath	150
commande.xml	150
Nœud ? Vous avez dit nœud ?	151
Chemins de localisation	151
Construire un chemin de localisation	152
Pourquoi XPath n'utilise-t-il pas la syntaxe XML ?	153

Chemins de localisation plus spécifiques	154
Fonctions XPath	155
Fonctions de nœud	156
Fonctions de position	158
Fonctions numériques	159
Fonctions booléennes	159
Fonctions de chaînes	161
Noms d'axes	165
Résumé	169

Chapitre 5 : XSLT en détail

Modèles	171
Ordre des opérations XSLT	172
Modèles par défaut	173
Comment un modèle affecte-t-il le nœud contextuel ?	173
Éléments XSLT et leur utilisation	174
<xsl:stylesheet>	174
<xsl:template>	175
Priorité de modèle	176
<xsl:apply-templates>	176
<xsl:value-of>	179
<xsl:output>	181
<xsl:element>	184
<xsl:attribute> et <xsl:attribute-set>	187
Groupes d'attributs apparentés	187
<xsl:text>	190
Traitement conditionnel avec <xsl:if> et <xsl:choose>	191
<xsl:for-each>	196
<xsl:copy-of>	199
<xsl:copy>	201
<xsl:sort>	204
Modes	206
Variables, constantes et modèles nommés	210
<xsl:variable>	211
Modèle nommé	213
Paramètres	213
Utiliser XSLT avec CSS	217
CSS et XSL-Formatting Objects	222
XSLT et le monde réel	222
Résumé	223

Chapitre 6 : Le modèle objet de document (DOM)

Qu'est-ce que le DOM ?	226
XML est un modèle objet	226
Modèle objet de document de XML	227

DOMString	228
Interfaces du DOM	228
Que sont ces interfaces ?	229
Implémenter des interfaces	229
Implémentations du DOM	231
DOM Core	232
Étendre les interfaces	232
DOM Core en détail	233
Exceptions	234
DOMException	235
Node	235
Extraire les informations d'un nœud	235
Naviguer dans l'arbre	238
Ajouter et supprimer des nœuds	240
Document	243
DOMImplementation	246
DocumentFragment	247
NodeList	248
Element	250
NamedNodeMap	253
Attr	255
CharacterData et Text	256
Gérer les chaînes complètes	257
Gérer les sous-chaînes	257
Modifier des chaînes	257
Fractionner du texte	258
Comment et CDATASection	262
ProcessingInstruction	263
Résumé	266

Chapitre 4 : XSLT et XPATH

Chapitre 7 : L'API simplifiée de XML (SAX)	267
---	-----

Vue d'ensemble	267
Le fabuleux trésor enfoui des Xenics	268
Et SAX dans tout ça ?	268
Télécharger SAX	269
Qui est donc David Megginson ?	270
Utiliser les interfaces SAX	270
Installation et configuration	270
Recevoir des événements SAX	274
Extraire des données textuelles	279
Extraire des attributs	283
Gestion des erreurs	287
Gestion des erreurs (suite) : objet Locator	290
Gestion des erreurs (suite et fin) : intercepter les erreurs d'analyse	292
Autres méthodes de DocumentHandler	295
ignorableWhitespace	295

processingInstruction	295
Qualités et défauts de SAX	296
SAX 2.0	297
Résumé	297

Chapitre 8 : Espaces de noms

À quoi servent les espaces de noms ?	299
Préfixes	301
Pourquoi opter pour une autre méthode ?	303
Espaces de noms XML	304
Espaces de noms par défaut	307
Espaces de noms des descendants	308
Annuler les espaces de noms par défaut	310
Notations différentes, résultats différents ?	311
Espaces de noms et attributs	313
XSLT et les espaces de noms	315
URI	319
URL	320
URN	320
Choisir entre URL et URN	321
Signification des URI	322
Quand doit-on utiliser les espaces de noms ?	323
Résumé	325

Chapitre 9 : Document XML valide : les DTD

Vocation et atouts des DTD	327
Ressources requises	328
Définitions de type de document	328
Validité de XML	329
Parseurs validateurs	330
Partage de DTD	330
Un exemple simple	332
Pourquoi utiliser XML ?	333
Modèle de données CatalogueLivres de base	333
DTD	338
DTD : internes ou externes ?	338
Quand les utiliser ?	339
Associer une DTD à un document XML	340
Déclaration XML	340
Déclaration DOCTYPE	341
Balises de base des DTD	346
Déclarations de type d'élément (ELEMENT)	347
Catégories de contenu d'élément	347
Modèles de contenu	349
Déclarations ATTLIST	358

Types d'attribut	360
Valeurs par défaut des attributs	365
Enrichir le modèle de données CatalogueLivres	367
Limites des DTD	381
Résumé	384

Chapitre 10 : Document XML valide : les schémas 387

Schémas et DTD	387
Syntaxe	388
Modèles de contenu	388
Typage des données	389
Extensibilité	389
Schémas dynamiques	390
Principes de base relatifs aux schémas	391
Parseurs XML et support des schémas	391
Structures et types de données, ou l'histoire de l'œuf et de la poule	392
Types de données	392
Types de données primitifs	393
Types primitifs de XML Schema	393
Types de données dérivés	394
Types dérivés intégrés des schémas XML	395
Types de données atomique et liste	395
Types de données atomiques	395
Types de données liste	396
Facettes des types de données	396
Espaces de valeurs	397
Espaces lexicaux	397
Facettes	398
Facettes fondamentales	398
Facettes de contrainte	400
Associer des schémas à des documents XML	403
Préambule d'un schéma XML : élément <schema>	403
Spécifier les espaces de noms	403
Spécifier la version	404
Préambule XDR : élément <Schema>	404
Structures	405
Définitions de type de données	405
Définitions de type simple	405
Définitions de type complexe	408
Modèles de contenu	414
Catégories de contenu d'élément	415
Cardinalité : minOccurs et maxOccurs	415
Groupes de choix et de séquences : <choice> et <sequence>	416
Contenu mixte (groupe non ordonné) : <all>	417
Contenu en provenance d'un autre schéma : <any>	418
Groupes de modèles : <group>	418

Déclarations d'élément	419
Contraintes d'élément : <element>	420
Déclarations d'attribut	422
Contraintes d'attribut : <attribute>	422
Groupes d'attributs : <attributeGroup>	424
Annotations	426
Résumé	427

Chapitre 11 : Document XML valide avancé 429

Entités	430
Références d'entités	431
Références de caractères	432
Entités générales	433
Entités analysables	433
Entités non analysables	438
Entités paramètres	438
Normes ISO et autres sur les entités	441
Données non XML (notations)	443
Déclaration <!NOTATION>	444
Type d'attribut NOTATION	444
Types d'attribut ENTITY et ENTITIES	444
Sous-ensembles de DTD internes ou externes	447
Attribut standalone	448
Sections conditionnelles	448
Imbriquer des sections conditionnelles	452
Résumé	452

Chapitre 12 : Les liens XML 453

Liens HTML	454
Liens XML	455
XLink	457
Attributs XLink	457
Attribut type	458
Attribut href	459
Attributs sémantiques (role et title)	459
Attributs comportementaux (actuate et show)	460
Attributs from et to	461
Types de lien	462
Liens simples	462
Liens étendus	465
Éléments du type locator (localisateur)	466
Arcs	467
Éléments du type resource (ressource)	469
Éléments du type title (titre)	471
Affecter une valeur par défaut aux attributs XLink	472

Pointer vers des parties de document à l'aide de XPointer	472
Ajouter des expressions XPointer aux URI	473
Utiliser plusieurs expressions XPointer	474
Schémas XPointer	476
Syntaxes XPointer raccourcies	476
Syntaxe complète	477
Syntaxe des noms abrégés	477
Syntaxe de la séquence des enfants	477
Positions, points et régions	478
Points	479
Régions	480
Comment sélectionner des régions ?	480
Fonctions d'extension XPointer à XPath	483
Requêtes	483
Résumé	484

Chapitre 13 : XML et les bases de données 485

Bases de données : d'hier à aujourd'hui	486
SQL	487
Extraire des données à partir de plusieurs tables	488
Normalisation	489
Utiliser vos propres applications	490
Architecture multiniveau	490
Utiliser XML dans une application multiniveau	492
Renvoyer des documents XML à partir d'un objet données	494
Intégrer XML dans la base de données	496
Stocker XML dans une base de données	496
Utiliser la base de données	497
Un objet données encore plus simple	498
Éditeurs de SGBD et XML	498
Technologies XML de Microsoft	498
MSXML	499
Générateur de code Visual Basic	499
SQL Server	499
Exécuter SQL Server	500
Technologies XML d'Oracle	502
Parseurs XML	502
Générateurs de code	503
Utilitaire XML SQL pour Java	503
Servlet XSQL	504
Exécuter Oracle	504
Au fait, qui a besoin d'une base de données ?	505
Résumé	506

Chapitre 14 : Autres utilisations de XML

507

Sérialiser les modèles objet

508

Suggestions pour obtenir un modèle objet efficace

515

Inclure des caractères de mise en forme

515

Autoriser l'auto-sérialisation de chaque objet

516

Fonction de validation/annulation

524

Ne rien renvoyer d'un objet vide

525

Suivre l'historique et les erreurs

526

XML-RPC

527

SOAP

529

RPC du pauvre

529

Objets sans état

535

Format de définition de ressources (RDF)

537

Fonctionnement

538

Schematron

539

Résumé

543

Étude de cas 1 : La boutique électronique :

les écrous de Sophie

545

Complexité

546

Mise à jour

548

Et l'ordre naquit du chaos

549

Concevoir la structure XML

549

Créer des fichiers XML

550

Regrouper les produits

554

Transformer et afficher XML avec XSL

554

Page web existante

557

Feuille de style XSLT

559

Incorporer un contenu à partir de sources multiples

567

Listes de produits

568

Modèles de « Top Écrous »

571

Modèles de « Au petit boulon »

571

Nouvelles pages web HTML

572

Avenir des écrous de Sophie

574

Étude de cas 2 : Un système de groupe de discussion	575
Concevoir l'architecture	576
XHTML	577
Remarques sur le canevas XMLPipes	577
Qu'accomplit le canevas XHTML ?	578
Définir le serveur	581
Définir les annuaires d'adhésion	582
Concevoir les panneaux d'affichage de messages	585
Récupérer l'arbre des enfants d'un message	589
Récupérer l'arbre des ancêtres d'un message	590
Récupérer les frères d'un message : message suivant	590
Récupérer les frères d'un message : message précédent	591
Créer des messages	592
Utiliser les espaces de noms	593
Modifier et supprimer un message	600
Supprimer un message	602
Afficher une liste de messages	603
À propos des éditeurs en ligne	609
Modifier et visualiser les messages individuels	610
Résumé	619
 Étude de cas 3 : XML pour une application interentreprises (B2B)	 621
Un problème de commerce électronique et notre solution	622
Exigences du commerce électronique	624
Pourquoi un EDI ne peut-il y arriver ?	625
Concept eBIS-XML	626
eBIS-XML, qu'est-ce que c'est précisément ?	629
Application eBIS-XML existante	630
Décisions, décisions...	631
DTD et schéma	631
Feuilles de style et XSL	632
Outils de développement de schémas	633
Schémas eBIS-XML	636
Développer les schémas eBIS-XML version 2.4	636
Un pas en avant : version 3	640
Schéma final	646
Résumé des astuces de développement du schéma	649
Retour sur les outils utilisés	649
Feuilles de style eBIS-XML	654
Une facture simple	654
Modèles poussés et tirés	658
Plus de détails sur la feuille de style de la facture	659
Choisir la voie facile ?	669

Résumé de la feuille de style	670
Avons-nous répondu aux exigences d'Arthur ?	670
Étapes suivantes	672
Routage simple	673
En-tête BizTalk	673
Chorégraphie	675
Autres possibilités	675
Et s'il fallait commencer dès maintenant ?	676
Résumé	677

Annexe A : Le modèle objet de document

Interfaces fondamentales	679
Interface DOMException	679
Interface Node	681
Interface Document	683
Interface DOMImplementation	685
Interface DocumentFragment	685
Interface NodeList	685
Interface Element	686
Interface NamedNodeMap	687
Interface Attr	688
Interface CharacterData	688
Interface Text	689
Interface Comment	689
Interfaces étendues	689
Interface CDATASection	690
Interface ProcessingInstruction	690
Interface DocumentType	690
Interface Notation	691
Interface Entity	691
Interface EntityReference	691

Annexe B : Une API simplifiée pour XML

Hiérarchie de classe	693
Interface org.xml.sax.AttributeList	694
Interface org.xml.sax.DocumentHandler	694
Interface org.xml.sax.DTDHandler	697
Interface org.xml.sax.EntityResolver	701
Interface org.xml.sax.ErrorHandler	703
Interface org.xml.sax.HandlerBase	705
Classe org.xml.sax.InputSource	707
Interface org.xml.sax.Locator	709
Interface org.xml.sax.Parser	713
Classe org.xml.sax.SAXException	714
Classe org.xml.sax.SAXParseException	717

Annexe C : Tutoriel d'apprentissage rapide ASP	721
Anatomie du protocole HTTP	722
Vision générale	722
Serveur HTTP	722
Notions de base sur les protocoles	723
Connexion	723
Requête	724
Réponse	727
Déconnexion	727
Introduction à Active Server Pages	728
Comment le serveur reconnaît les pages ASP	728
Notions de base sur ASP	729
Balises ASP	729
Blocs <SCRIPT>	729
Langage de script par défaut	729
Combiner HTML et ASP	730
Commenter votre code ASP	731
Modèle d'objets ASP	732
Collections	733
Parcourir la collection Contents	733
Supprimer un élément de la collection Contents	734
Objet Request	735
Collections de l'objet Request	735
Propriétés et méthodes de l'objet Request	739
Objet Response	739
Collection de l'objet Response	739
Méthodes de l'objet Response	741
Propriétés de l'objet Response	742
Objets Application et Session	744
La portée est éternelle	744
Fichier global.asa	744
Objet Application	744
Objet Session	746
Objet Server	748
Propriété	748
Méthodes	748
ObjetObjectContext	750
SetAbort	750
SetComplete	750
Utiliser Active Server Pages de façon efficace	750
Concevoir le site	750
Créer le fichier global.asa	751
Créer la page principale	752
Section ASP/VBScript	752
Section HTML	754
Résumé	756

Annexe D : Tutoriel d'apprentissage rapide JavaScript	757
JavaScript et HTML	758
JavaScript et Java	758
Exécuter les exemples de cette annexe	759
Instructions	760
Commentaires	761
Commentaires sur une ligne	761
Commentaires multilignes	761
Variables	762
Types de données	763
Noms de variables	764
Opérateurs	765
Opérateurs mathématiques	765
Opérateurs d'affectation	766
Opérateurs logiques (booléens)	766
Préséance des opérateurs	767
Fonctions	768
Créer une fonction	768
Renvoyer des informations à partir d'une fonction	768
Transmettre des informations à une fonction	769
Fonctions intégrées	769
Flux d'un programme	770
Code conditionnel	770
Instruction if	770
Instruction switch	772
Boucles	773
Boucle for	773
Boucles while and do ... while	774
Instructions break et continue	775
Objets	776
String	776
Méthodes toLowerCase et toUpperCase	777
Méthodes renvoyant des chaînes et caractères	777
Méthodes de recherche dans une chaîne	778
Concaténer des chaînes	778
Number	778
Math	779
Date	779
Array	780
Résumé	780

Annexe E : Ressources web utiles 781

Annexe F : Types de schémas XML 786

Types primitifs 787

string	788
boolean	788
decimal	788
float	789
double	789
timeDuration	790
recurringDuration	791
binary	792
uriReference	792
ID	792
IDREF	792
ENTITY	793
NOTATION	793
QName	793

Facettes de contrainte des types primitifs 794

Types dérivés intégrés 795

language	796
name, NCName	796
integer, negativeInteger, positiveInteger, nonNegativeInteger, nonPositiveInteger	796
byte, short, int, long	796
unsignedByte, unsignedShort, unsignedInt, unsignedLong	797
century, year, month, date	797
recurringDate, recurringDay	797
time, timeInstant, timePeriod	798
Facettes de contrainte des types dérivés	798

Index 801