

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

THESE

Présentée devant

L'INSTITUT NATIONAL DE FORMATION EN INFORMATIQUE (ex : CERI) - ALGER

Pour obtenir

LE GRADE DE DOCTEUR D'ETAT ÈS SCIENCES

Discipline : Informatique

Par

Azeddine CHIKH

(Ingénieur CERI)

**Une méthodologie de réutilisation
en ingénierie du document
Le système SABRA**

Soutenue le 20 Juillet 2004, devant le jury composé de :

H. Drias,	Professeur, INI - Alger	Algérie	Président
Z. Boufaïda,	Professeur, Université de Constantine	Algérie	Examineur
M. Ahmed Nacer,	Professeur, USTHB	Algérie	Examineur
M. Boufaïda,	Professeur, Université de Constantine	Algérie	Examineur
T. Laskri,	Professeur, Université de Annaba	Algérie	Examineur
M. Abbas,	Professeur, USTHB	Algérie	Co-Directeur
M-F. Barthet,	Professeur, Université de Toulouse I	France	Directeur de thèse

Thèse préparée au sein du laboratoire des Interactions Homme-Système LIHS – Université Toulouse I
et du laboratoire de Génie-BioMédical GBM – Université Abou Bakr Belkaid - Tlemcen

ملخص

لقد دخلنا عصر إنتاج الوثائق بواسطة تجميع الأجزاء الموجودة وهي مقاربة تتضمن فوائد بديهية: تحسين الإنتاجية، تطبيع اكبر، و تماسك أفضل بين الوثائق. إن إعادة الاستعمال في هندسة الوثيقة تعتبر محور بحث جديد. لذلك فهي تعاني على الخصوص من نقص في المنهجية. لهذا جاء هذا العمل ليهتم بهذا الميدان و يحاول أن يحدد طريقة منهجية تسمى ARBRE (طريقة تساعد على التحرير المستند إلى إعادة الاستعمال) لها القدرة على تحديده و التكفل به. إن هذه المنهجية تسمح بتعريف إعادة الاستعمال في هندسة الوثيقة و تتطرق إليه من خلال بعدين اثنين: سكوني و تحريكي: أما البعد الأول فيتعلق بالمفاهيم المعالجة في حين يتعلق الثاني بالأنشطة التي تتعلق به. كما نقترح نموذجين : يسمى الأول ASARD (الهوامش المهيكلة للمساعدة في إعادة استعمال الوثائق) و يتكفل بتحديد المعرفة القابلة لإعادة الاستعمال (اللبنات) و تعيين احتياجات جديدة في الوثائق الجديدة المعدة للا نجاز. أما النموذج الثاني و يسمى MCD (نموذج مكون الوثيقة) فيتكفل ببيان مكونات إعادة الاستعمال.

بالإضافة إلى هذا فإننا نقترح نظام SABRA (نظام للتحرير عن طريق إعادة الاستعمال يركز على الهوامش) و يمثل سندا تقنيا للمنهجية ARBRE. و في الأخير نوضح مدى تطبيقية نتائج بحثنا على حالة تعليمية تتصل بموضوع انجاز درس في الهندسة البرمجية.

RESUME

Nous sommes entrés dans l'ère de la production des documents par assemblage de fragments existants. Les bénéfices d'une telle approche sont évidents : amélioration de la productivité, plus grande normalisation et meilleure cohérence des documents, les uns avec les autres.

La réutilisation, en ingénierie du document, est un nouvel axe de recherche. L'état de l'art en la matière souffre essentiellement du manque de méthodologie. Ce travail de recherche s'intéresse à ce domaine et tente de définir une approche méthodologique, baptisée *ARBRE* (Approche d'aide à la **R**édaction **B**asée sur la **R**Eutilisation), capable de le cerner et de le prendre en charge. Cette méthodologie permet de définir la réutilisation en ingénierie du document et de l'aborder selon deux dimensions : statique et dynamique. La première est relative aux entités manipulées, la seconde est relative aux activités et acteurs impliqués. Nous proposons aussi deux modèles sous-jacents. Le premier, baptisé *ASARD* (Annotations Structurées d'Aide à la **R**éutilisation de **D**ocuments), prend en charge la qualification de la connaissance réutilisable (les briques) et la spécification des nouveaux besoins dans les nouveaux documents à construire. Le second, baptisé *MCD* (Modèle de Composant de **D**ocument), prend en charge la représentation des composants de réutilisation.

Nous proposons en outre le système *SABRA* (System of **A**uthoring **B**y **R**euse based on **A**nnotations), qui sert de support technique à la méthodologie *ARBRE*. Enfin nous montrons l'applicabilité de nos résultats de recherche sur un cas de e-learning portant sur la construction d'un cours sur le génie-logiciel.

ABSTRACT

We are entering the age of document production by assembling existant fragments. The benefit of such approach is obvious : enhancing of productivity, greater normalisation and better document coherence.

The reuse, in document engineering, is a new research orientation. The state of the art on the matter suffers primarily from the lack of methodology. This research work is interested in this field and tries to define a methodological approach, called *ARBRE*, able to deal with it. This approach makes it possible to define the reuse in document engineering and to describe it according to two dimensions : static and dynamic. The former relates to the handled objects, while the latter relates to the implied activities and actors. We propose also two subjacent models. The first, called *ASARD*, deals with the qualification of bricks of document. The second called *MCD* and including the first, deals with the representation of the document components.

We propose the system *SABRA*, which is used as technical support for the methodology *ARBRE*. We show the applicability of these research results on an e-learning case study related to constructing a new course of software engineering.

TABLE DES MATIERES

TABLE DES MATIERES

INTRODUCTION GÉNÉRALE	17
1. PROBLÉMATIQUE.....	18
2. CONTRIBUTIONS.....	18
3. PLAN DU RAPPORT.....	18
PARTIE 1 - DOMAINES D'INTÉRÊT	23
1. CHAPITRE – RÉUTILISATION DES CONNAISSANCES	25
1.1. <i>Réutilisation de connaissances et raisonnement analogique</i>	25
1.2. <i>Mémoire sémantique versus mémoire épisodique</i>	26
1.3. <i>Représentation de la réutilisation</i>	27
2. CHAPITRE - PRODUCTION ÉCRITE	29
2.1. <i>Modèle des stades</i>	29
2.2. <i>Modèle de résolution de problèmes</i>	29
2.3. <i>Modèle de l'interaction sociale</i>	31
3. CHAPITRE - INGÉNIERIE DU DOCUMENT.....	33
3.1. <i>Introduction</i>	33
3.2. <i>Document vs Document électronique</i>	33
3.2.1. Définition du document	33
3.2.2. Définition du document électronique et de la GED	33
3.3. <i>Processus de la GED</i>	34
3.4. <i>Cycle de vie d'un document</i>	35
3.4.1. Axe horizontal	35
3.4.2. Axe vertical	35
3.5. <i>Structuration de document</i>	36
3.5.1. Structure physique	37
3.5.2. Structure logique	38
3.5.3. Complémentarité des deux structures : logique et physique	39
3.5.4. Structure spécifique , structure générique et modèle de document	40
3.6. <i>Activités de l'ingénierie du document</i>	40
3.7. <i>Qualité du document</i>	41
3.8. <i>XML</i>	42
3.8.1. Pourquoi XML ?	42
3.8.2. Avantages de XML	42
3.8.3. Intérêt de XML pour la mémoire documentaire	43
3.8.4. Quelques technologies autour du langage XML	44
3.8.5. Applications existantes	44
4. CHAPITRE - MÉTA DONNÉES	47
4.1. <i>Définitions</i>	47
4.2. <i>Modèles de méta données</i>	47
4.2.1. Schémas spécifiques	47
4.2.2. Schémas standards	48
4.3. <i>Le modèle LOM</i>	49
5. CHAPITRE - DOCUMENTS ADAPTATIFS.....	51
5.1. <i>Introduction</i>	51
5.2. <i>Critères d'adaptation</i>	51
5.3. <i>Éléments d'adaptation</i>	52
5.4. <i>Présentation adaptative</i>	53
5.4.1. Méthodes de la présentation adaptative	53
5.4.2. Techniques de la présentation adaptative	54
5.5. <i>Support de navigation adaptative</i>	54
5.5.1. Méthodes du support de navigation adaptative	55
5.5.2. Techniques du support de navigation adaptative	56
6. CHAPITRE - DOCUMENTS VIRTUELS	59
6.1. <i>Introduction</i>	59
6.1.1. Structure à trois niveaux	59
6.1.2. Niveau central : le document comme structure abstraite	59
6.1.3. Niveau de surface : la scénation et la scénique comme formes d'inscriptions spatiales et temporelles	60
6.2. <i>Définitions des documents virtuels</i>	60

7.	CHAPITRE - MÉMOIRES DOCUMENTAIRES	61
7.1.	<i>Introduction</i>	61
7.2.	<i>Entrepôt de documents</i>	61
7.3.	<i>Processus de développement</i>	62
8.	CHAPITRE - ONTOLOGIES.....	63
8.1.	<i>Définitions</i>	63
8.2.	<i>Pourquoi des ontologies?</i>	63
8.3.	<i>Modélisation des ontologies</i>	64
8.4.	<i>Technologie RDF (Resource Description Framework)</i>	64
8.4.1.	<i>Introduction</i>	64
8.4.2.	<i>Utilisation</i>	65
PARTIE 2 - ETAT DE L'ART SUR LA RÉUTILISATION EN INGÉNIERIE DU DOCUMENT		69
9.	CHAPITRE – RÉFLEXIONS THÉORIQUES	71
9.1.	<i>Réutilisation de documents</i>	71
9.1.1.	<i>Pourquoi réutiliser en ingénierie du document ?</i>	71
9.1.2.	<i>Processus de production par réutilisation</i>	72
9.2.	<i>Réutilisation de matériaux pédagogiques</i>	72
9.2.1.	<i>Processus de préparation de matériaux pédagogiques «for reuse»</i>	72
9.2.2.	<i>Apports des nouvelles technologies à la réutilisation de matériaux pédagogiques</i>	73
9.3.	<i>Méthode de composition de documents pédagogiques adaptatifs</i>	73
9.4.	<i>Généralisation des structures de documents et assemblage de documents</i>	74
9.5.	<i>Approche modulaire, flexible et évolutive</i>	74
9.6.	<i>Approche de réutilisation basée sur la connaissance</i>	75
10.	CHAPITRE – MODÈLES DE LA CONNAISSANCE RÉUTILISABLE.....	77
10.1.	<i>Modèle de la base de connaissance</i>	77
10.2.	<i>Méta-structure documentaire</i>	78
10.3.	<i>Méta-modèle de document</i>	79
10.4.	<i>Modèle d'entrepôt de document</i>	80
11.	CHAPITRE – PROJETS DE RÉUTILISATION D'ITEMS DIDACTIQUES.....	81
12.	CHAPITRE – BILAN DE L'ÉTAT DE L'ART ET NOUVEAUX OBJECTIFS.....	85
12.1.	<i>Synthèse de l'état de l'art</i>	85
12.2.	<i>Limites des travaux présentés et objectifs de notre travail</i>	86
PARTIE 3 - RÉSULTATS DE NOTRE RECHERCHE		91
13.	CHAPITRE – MONDE DE SUJET : LA MÉTHODOLOGIE ARBRE	93
13.1.	<i>Introduction</i>	93
13.2.	<i>Définition de la réutilisation en ingénierie du document</i>	93
13.3.	<i>Dimension statique</i>	94
13.3.1.	<i>P_documents</i>	94
13.3.2.	<i>O_documents</i>	95
13.3.3.	<i>Briques</i>	96
13.3.4.	<i>Q_composants</i>	100
13.3.5.	<i>S_composants</i>	102
13.3.6.	<i>X_entrepôts</i>	103
13.3.7.	<i>X_bibliothèques</i>	105
13.4.	<i>Dimension dynamique</i>	106
13.4.1.	<i>Niveaux de maturité dans le processus de réutilisation</i>	106
13.4.2.	<i>Intégration de la réutilisation en ingénierie du document</i>	107
13.4.3.	<i>«Authoring for reuse» et «Authoring by reuse»</i>	107
13.4.4.	<i>Acteurs de la réutilisation</i>	108
13.4.5.	<i>Cycle de vie de production spécifique à la réutilisation</i>	109
14.	CHAPITRE - MONDE DE SYSTÈME : LE MODÈLE ASARD	113
14.1.	<i>Introduction</i>	113
14.2.	<i>Modèle général d'annotation structurée</i>	113
14.3.	<i>Qualification des briques dans le modèle ASARD</i>	116
14.3.1.	<i>Qualification par les méta données</i>	117
14.3.2.	<i>Qualification par les ontologies</i>	118
14.3.3.	<i>Qualification par les associations</i>	119
14.3.4.	<i>Qualification par le texte libre</i>	120
14.4.	<i>Spécification du o_document dans le modèle ASARD</i>	121
14.4.1.	<i>Spécification par les méta données</i>	122
14.4.2.	<i>Spécification par les ontologies</i>	123

14.4.3.	Spécification par les associations	123
14.4.4.	Spécification par le texte libre	124
15.	CHAPITRE – MONDE DE SYSTÈME : LE MODÈLE MCD	125
15.1.	<i>Représentation UML (Uniform Modeling Language)</i>	<i>125</i>
15.1.1.	Ilot de production	125
15.1.2.	Ilot de la connaissance réutilisable	126
15.1.3.	Ilot de la connaissance de qualification et de spécification	126
15.1.4.	Ilot d'organisation	127
15.2.	<i>Représentation XML (eXtensible Markup Language)</i>	<i>129</i>
15.2.1.	B_entrepôt	130
15.2.2.	Q_bibliothèque	135
15.2.3.	O_entrepôt	137
15.2.4.	S_bibliothèque	142
15.2.5.	D_entrepôt	144
15.2.6.	Catalogues	147
15.2.7.	Hypérisation du b_entrepôt	147
15.3.	<i>Représentation OEM (Object Exchange Model)</i>	<i>151</i>
16.	CHAPITRE – LE SYSTÈME SABRA	157
16.1.	<i>Introduction</i>	<i>157</i>
16.2.	<i>Architecture</i>	<i>157</i>
16.2.1.	Module de base	159
16.2.2.	Module “Authoring for reuse”	159
16.2.3.	Module “Authoring by reuse”	160
16.3.	<i>Monde de développement : «Authoring for reuse»</i>	<i>161</i>
16.4.	<i>Monde d’usage : «Authoring by reuse»</i>	<i>162</i>
16.5.	<i>Implémentation d’un prototype</i>	<i>163</i>
16.5.1.	Module de base	163
16.5.2.	Module “Authoring for reuse”	167
16.5.3.	Module “Authoring by reuse”	170
	PARTIE 4 – CAS D'APPLICATION : E-LEARNING	179
1.	INTRODUCTION.....	179
2.	BESOINS EN RÉUTILISATION DANS LE DOMAINE DU E-LEARNING.....	179
3.	APPLICATION.....	180
	CONCLUSION	189
	RÉFÉRENCES BIBLIOGRAPHIQUES	193
	GLOSSAIRE DES ABRÉVIATIONS UTILISÉES	201

TABLE DES FIGURES

Figure 1. Domaines d'intérêt	23
Figure 2. Représentation cognitive de la réutilisation.....	27
Figure 3. Modèle de processus d'écriture selon Flower & Hayes (extraite de [FLOW81]).....	30
Figure 4. Cycle de vie du document (extraite de [PARE01])	36
Figure 5. Exemple de structure physique d'un document (extraite de [QUIN94])	38
Figure 6. Exemple de structure logique d'un document (extraite de [QUIN94]).....	39
Figure 7. Méthodes et techniques d'adaptation des hypermédias	52
Figure 8. Modèle de base de connaissance	77
Figure 9. Méta-structure documentaire	78
Figure 10. Méta-modèle de document.....	79
Figure 11. Schéma UML de l'entrepôt [KHRO01].....	80
Figure 12. Organisation bidimensionnelle des briques	100
Figure 13. Schéma général de la réutilisation en ingénierie du document.....	109
Figure 14. Cycle de production spécifique à la réutilisation.....	110
Figure 15. Schéma général d'une annotation du modèle ASARD.....	113
Figure 16. Catégories d'annotations dans le modèle ASARD	114
Figure 17. <i>Q_annotations</i> du modèle ASARD	116
Figure 18. Qualification des briques dans le modèle ASARD.....	116
Figure 19. <i>S_annotations</i> du modèle ASARD	121
Figure 20. Spécification du <i>o_document</i> dans le modèle ASARD	122
Figure 21. Modèle général de composant de document (UML)	128
Figure 22. Modèle OEM des briques	152
Figure 23. Modèle OEM des <i>q_annotations</i>	153
Figure 24. Modèle OEM des <i>q_composants</i>	154
Figure 25. Modèle OEM des <i>o_documents</i>	154
Figure 26. Modèle OEM des <i>s_composants</i>	155
Figure 27. Modèle OEM des <i>p_documents</i>	155
Figure 28. Architecture générale du système SABRA.....	158
Figure 29. Processus relatif à l'activité «Authoring for reuse».....	161
Figure 30. Processus relatif à l'activité «Authoring by reuse»	162
Figure 31. Etape 1 du processus d'indexation (choix de la technique).....	163
Figure 32. Etape 2 du processus d'indexation (création de l'index).....	164
Figure 33. Etape 3 du processus d'indexation (indexation en cours).....	164
Figure 34. Résultat de l'indexation (Propriétés)	165
Figure 35. Choix de techniques de stockage.....	165
Figure 36. Base de données générée avec l'approche EDGE	166
Figure 37. Structure de la table «Edge table»	166
Figure 38. Authoring for reuse : interface principale (capitalisation d'une brique)	167
Figure 39. Menu contextuel d'une brique	168
Figure 40. Authoring for reuse : étape 1 de la qualification par ASARD.....	168
Figure 41. Authoring for reuse : étape 2 de la qualification par ASARD.....	169
Figure 42. Authoring for reuse : <i>q_annotations</i> existantes	169
Figure 43. Authoring by reuse : interface principale (menu Projet)	170
Figure 44. Authoring by reuse : interface principale (menu Outline).....	171
Figure 45. Authoring by reuse : interface principale (menu Document)	171
Figure 46. Authoring by reuse : interface principale (menu Outils).....	172
Figure 47. Authoring by reuse : interface de document outline (spécification LOM).....	172

Figure 48. Authoring by reuse : interface de document outline (briques résultats)	173
Figure 49. Authoring by reuse : interface de document outline (navigation)	174
Figure 50. Authoring by reuse : interface de document outline (gestion).....	175
Figure 51. Authoring by reuse : interface de document à construire (onglet Document).....	175

TABLE DES LISTINGS

Listing 1. DTD des briques (bricks.dtd).....	130
Listing 2. Instance XML des briques (bricks.xml).....	131
Listing 3. DTD des <i>q_annotatiions</i> (q_asard.dtd).....	133
Listing 4. Instance XML des <i>q_annotatiions</i> (q_asard.xml).....	134
Listing 5. DTD des <i>q_composants</i> (q_library.dtd).....	135
Listing 6. Instance XML des <i>q_composants</i> (q_library.xml).....	136
Listing 7. DTD des <i>o_documents</i> (o_house.dtd).....	137
Listing 8. Instance XML des <i>o_documents</i> (o_house.xml).....	138
Listing 9. DTD des <i>s_annotatiions</i> (s_asard.dtd).....	140
Listing 10. Instance XML des <i>s_annotatiions</i> (s_asard.xml)	141
Listing 11. DTD des <i>s_composants</i> (s_library.dtd).....	142
Listing 12. Instance XML des <i>s_composants</i> (s_library.xml)	143
Listing 13. DTD des <i>p_documents</i> (d_house.dtd).....	145
Listing 14. Instance XML des <i>p_documents</i> (d_house.xml).....	146
Listing 15. DTD des briques - version hyperdocument (bricks.dtd).....	148
Listing 16. DTD relative à la structure hypertexte du <i>b_entrepôt</i> (pivot_b_house.dtd)	149
Listing 17. Instance XML des briques - version hyperdocument (bricks.xml).....	149
Listing 18. Instance XML de la structure hypertexte du <i>b_entrepôt</i> (pivot_b_house.xml) ...	150
Listing 19. Instance XML du <i>o_document</i> relatif au cours de GL à construire.....	180
Listing 20. Instance XML des <i>s_annotatiions</i> relatives au <i>o_élément</i> 'e31'	181
Listing 21. Requête Xquery relative aux <i>s_annotatiions</i> Meta_describe du listing 20.....	182
Listing 22. Contenu de la brique relative au modèle de cascade (b100.xml)	182
Listing 23. Contenu de la brique relative au modèle en spirale (b105.xml)	183
Listing 24. Contenu de la brique relative au modèle en V (b110.pdf).....	183
Listing 25. Contenu de la brique relative au modèle par incrément (b123.ps)	183
Listing 26. Instance XML des briques pédagogiques à réutiliser dans le cours de GL	183
Listing 27. Instance XML de <i>q_annotatiions</i> relatives à la brique 'b100'.....	184
Listing 28. Instance XML du <i>p_document</i> relatif au cours de GL.....	185
Listing 29. Présentation finale du <i>p_document</i> avec une feuille de style	186

TABLE DES TABLEAUX

Tableau 1. Catégories de descripteurs du modèle LOM	49
Tableau 2. Correspondance : Brique - <i>P_document</i>	97
Tableau 3. Interprétation des flèches de la figure 13	109