

Partage des modèles XML : une solution pour les échanges électroniques professionnels

A. MKADMI, N. BOUHAI

Laboratoire PARAGRAPHÉ, Université Paris8,

2, rue de la Liberté

93526 - SAINT-DENIS cedex 02,

Mail : mkadmi@hymedia.univ-paris8.fr ; bouhai.nacer@free.fr

Tél : +33 1 49 40 67 58 Fax : +33 1 49 40 67 83

M. LANGLOIS

EDIFRANCE

13, rue Camille Desmoulins

92441 Issy-les-Moulineaux Cedex

Mail : langloismarc@compuserve.com

Tél : +33 1 58 04 25 32 Fax : +33 1 58 04 23 00

Résumé

Nous présentons dans cet article le projet de création d'un répertoire de modèles XML que nous avons réalisé dans le cadre d'une coopération entre le laboratoire Paragraphe, Edifrance, Mutu-XML, GFII, FING et UIC. Nous présentons également le contexte et l'intérêt général du projet. Une description de l'application avec ses différentes fonctionnalités concernant l'accès, la recherche, le partage et la révision des modèles sera présentée. Enfin, nous présenterons les résultats et perspectives de ce projet.

Abstract

We present in this article the project of creation of repository of XML models which we carried out within the framework of co-operation between laboratory Paragraphe, Edifrance, Mutu-XML, GFII, FING and UIC. We present also general context and interest of the projet. A description of the application with its various functionalities relating to access, research, division and revision of the models will be presented. Lastly, we will present the results and perspective of this project.

1 Introduction

Comme toute technologie émergente, il aura fallu attendre plusieurs années avant que le format XML (eXtensible Markup Language) ne s'impose réellement comme étant le format privilégié de l'échange de documents - et plus généralement d'informations - en milieu ouvert. En dissociant le contenu de la publication, XML est apparu en réponse au besoin d'interaction et de coopération entre des systèmes d'information hétérogènes utilisant jusqu'alors des structures de données largement incompatibles entre elles. En effet, les travaux de normalisation menés par le W3C, les développements des éditeurs de logiciels et les préconisations de différents groupes et consortiums (ebXML, OASIS) conjuguent leurs efforts pour définir, promouvoir et utiliser XML dans différentes situations. Ce méta-langage est utilisé aujourd'hui par tous : les fournisseurs d'ERP, les éditeurs de middleware, les fournisseurs de bases de données, etc. Les raisons de ce consensus sont à chercher du côté de la simplicité et de la richesse d'expression d'XML.

Cependant, pour que les objectifs d'XML, en l'occurrence permettre l'échange généralisé intersectoriel quel que soit le type d'acteur, soient réellement atteints, il paraît très intéressant que les modèles de ces documents structurés échangés, ainsi que toutes les informations associées soient partagés. Ces informations associées servent à mieux comprendre les modèles pour pouvoir les exploiter de façon rapide et efficace. Il devient alors possible de recevoir n'importe quel document issu d'un modèle particulier afin d'être en mesure de l'exploiter avec des logiciels génériques. On pourra, par exemple, éditer un document avec des environnements standard et disponibles, mais aussi le visualiser sur un système de consultation standard du Web.

2 Problématique

Associant aux données une structure sémantique (sous forme d'éléments et attributs) et permettant de séparer cette structure du contenu, ainsi que la présentation de ce contenu, XML a été retenu comme le langage d'avenir pour la génération des échanges électroniques, que ce soit entre les grandes entreprises, entre les PME ou entre les grandes entreprises et les PME. Comparé à l'EDI conventionnel, XML offre la capacité de :

- afficher les données sous une forme humaine par l'utilisation des feuilles de style... ;
- convertir facilement une structure de message en une autre structure de message (ce qui facilite l'intégration des données dans des applications existantes...) [GENCOD EAN, 2002]

Cependant, pour permettre un traitement automatique de documents XML provenant d'autres partenaires, il est nécessaire que les différents acteurs impliqués se mettent d'accord sur un formalisme de structuration des informations. Cette structuration est définie selon un modèle sous formats de schémas ou de DTD qui donne les règles d'assemblage et d'ordonnement des données.

De ce fait, il devient alors intéressant de pouvoir partager ces modèles. C'est dans ce contexte que notre projet de création d'un répertoire de modèles XML a vu le jour pour pouvoir identifier, partager et réutiliser les modèles de documents pour les différentes applications XML.

3 Partage des modèles de documents : état de l'art

Le partage et la réutilisation des modèles de documents étaient depuis longtemps le souci des professionnels, et si le déploiement de l'Internet décuple les possibilités de partage

d'information, les syntaxes utilisées jusqu'alors (HTML notamment) limitent les types de réutilisation possibles. Le développement de XML rend possible le partage de données structurées. Ceci correspond à une mutation profonde de la façon dont les applications informatiques vont partager des référentiels sémantiques, avec des conséquences importantes en terme d'infrastructures et d'outils disponibles. Nous présentons en ce qui suit quelques exemples de modèles de documents partagés utilisés dans le domaine public.

3-1 Répertoire de l'OASIS¹

Le répertoire de l'OASIS (OASIS registry/repository) est un système ayant, comme son nom anglais l'indique, deux composants, un « registre » et un « dépôt ». On parle donc des objets enregistrés et des entrées de registre. Les objets enregistrés représentent tout ce qu'un auteur met à la disposition du public pour être utilisé par un client. Ils sont sauvegardés dans le « dépôt ». Quant aux entrées de registre, elles représentent les métadonnées (informations décrivant les objets enregistrés), elles sont stockées dans le registre et servent de référence pour accéder aux objets.

Ce répertoire ouvert pour XML et SGML DTDs et schémas représente un pas pour déployer des répertoires XML interopérables sur Internet. Il offre un forum indépendant des éditeurs pour les programmeurs et organismes de standardisation pour soumettre publiquement, publier et échanger des spécifications et vocabulaires XML. Il est conçu pour servir de modèle pour un réseau extensible de registres et répertoires XML distribués sur Internet. Les spécifications du registre ebXML de l'OASIS sont disponibles à <http://www.oasis-open.org/committees/regrep/documents/2.0/specs>.

Cependant ce répertoire, malgré l'intérêt qu'il présente dans l'orientation vers une plus grande utilisation du langage XML, son utilisation n'est pas à la portée de tout le monde (qui n'est pas forcément spécialiste en XML). De plus son orientation anglophone peut être une contrainte pour d'autres acteurs (notamment francophones) au niveau de son utilisation.

3-2 Répertoire de l'ATICA²

Dans le domaine de l'administration, il existe aussi un projet similaire mené par l'ATICA visant à publier tous les schémas et DTD issus des domaines documentaire et juridique présentant un intérêt général dans un répertoire. Ce répertoire représente un outil mutualisé pour favoriser les échanges au sein des administrations et avec leurs partenaires.

Ce projet a été déclenché par une Circulaire du 21 janvier 2002 relative à la mise en œuvre d'un cadre commun d'interopérabilité pour les échanges et la compatibilité des systèmes d'information des administrations.

« Enfin, il sera bon que chaque nouveau projet de système comportant des échanges d'informations (au sein de l'administration ou avec les tiers) soit l'occasion de poursuivre, et même d'intensifier, l'élaboration de schémas XML, dont on connaît l'importance pour faciliter les échanges. Ils seront conçus de manière à faire clairement apparaître leur définition, ainsi que celle des éléments qui les composent, par application de la méthode dite des « espaces nominatifs », conforme aux standards de l'Internet. Ils seront publiés, d'abord à l'état de projet, puis sous leur forme définitive, dans le répertoire des schémas XML de l'administration. » [Jospin, 2002].

¹ OASIS Registry/Repository : technical Specification, Working Draft 1.1 december 20, 2000/ OASIS-Organization for the Advancement of Structured Information Systems.

² ATICA : Agence pour les Technologies de l'Information et de la Communication dans l'Administration

Cependant, ce projet, malgré l'intérêt qu'il présente quant à la dématérialisation et l'interopérabilité des échanges au sein de l'administration et entre l'administration et ses partenaires, il n'est pas précisé jusqu'à aujourd'hui comment les modèles seront présentés au sein du répertoire, ni quelles sont les documentations qui doivent accompagner ces modèles. Ces lacunes n'ont pas encouragé les administrations à se lancer dans ce projet pour enrichir le répertoire par leurs modèles.

Ces deux exemples de projets (qui ne sont pas exclusifs) montrent bien l'intérêt de la création d'un répertoire de schémas XML. En effet, Le fait de partager des modèles participe à l'évidence à l'acceptation de ceux-ci et représente, de plus, un facteur de montée en compétences des organisations confrontées à l'explosion des applications XML. Plus les modèles sont accessibles de manière facile, plus l'échange entre partenaires devient aisé.

4 Intérêt des modèles de documents

Un modèle de document est une structure permettant de donner les règles d'assemblage des données. Cette structure permet à un système d'information de comprendre que `<Code_Postal>96100</Code_Postal>` représente bien un code postal d'une ville et non pas une quelconque suite de chiffres. Ce système peut aussi comprendre, d'un point de vue structurelle, que ce code fait partie d'une adresse d'un client. À défaut de tel modèle, ce numéro peut être codé de différentes manières dans des différents documents, exemple : `<codepostal>`, `<CodePostal>`, ..., et par conséquent, il peut être interprété comme une suite de chiffres.

Un modèle de documents sert donc dans un premier temps à définir tous les éléments utilisés dans un document, et deuxièmement pour définir les relations et l'ordonnancement entre ces éléments.

Pour créer des modèles de documents, deux recommandations existent aujourd'hui. Les DTDs et les schémas XML. La première sur les DTDs est historique et est devenue reconnue dans tous les domaines applicatifs. La deuxième est récente et vient palier aux déficiences de la première se rapportant notamment au typage de données, au langage utilisé et au support des espaces de noms. En effet, XML Schema est un nouveau langage proposé par le W3C qui propose, en plus des fonctionnalités fournies par les DTD, plusieurs nouveautés à savoir :

- un grand nombre de types de données intégrées comme les booléens, les entiers, les intervalles de temps, etc. De plus, il est possible de créer de nouveaux types par ajout de contraintes sur un type existant ;
- des types de données utilisateurs qui nous permettent de créer notre propre type de données nommé ;
- la notion d'héritage : Les éléments peuvent hériter du contenu et des attributs d'un autre élément. C'est sans aucun doute l'innovation la plus intéressante de XML Schema ;
- le support des espaces de nom ;
- les indicateurs d'occurrences des éléments peuvent être tout nombre non négatif ;
- une grande facilité de conception modulaire de schémas.

Les modèles de documents servent donc à définir la cohérence d'un ensemble de documents, lesquels peuvent être utilisés par n'importe quelle application informatique en ne se définissant que par rapport au modèle sous-tendu. Ceci permet évidemment de gagner beaucoup de temps, d'argent et de fiabilité dans les travaux coopératifs.

5 Présentation du répertoire de modèles XML

Le répertoire de modèles XML est une base de données permettant à tout utilisateur de prendre connaissance des modèles existants dans un domaine d'activité particulier pour un besoin particulier, ainsi que des modèles permettant d'échanger avec ses partenaires. Ce répertoire favorise l'échange ouvert entre professionnels qui, avec cet outil, seront capables de se définir par rapport à l'existant et d'avoir une cohérence des méthodes de travail dans leurs domaines d'activités.

Le répertoire de modèles XML doit alors proposer des accès à un ensemble de modèles bien documentés donnant une idée sur les initiatives normatives prises dans un domaine d'activité bien déterminé. Il doit représenter aussi un espace de travail collaboratif dans le cadre des échanges électroniques professionnels.

5-1 Principes de base

Le répertoire de modèles est accessible librement à tout le monde et il n'y a aucune limite dans la consultation des schémas et des DTDs qui y sont stockés. Quant à la sécurité, il n'y a pour le moment aucune notion de confidentialité au niveau accès des données. La recherche des modèles, comme nous allons la présenter par la suite, se fait à travers plusieurs critères définissant le contexte et l'appartenance de chaque modèle. Quant à la modification et la mise à jour, elles se font par le propriétaire des modèles, seuls responsables des structures de leurs documents.

Cependant, pour la publication de ces modèles, elle se passe par un « comité éditorial » qui vérifie la forme de ce qui est proposé à publication, c'est à dire la cohérence des données contextuelles, la pertinence de la définition par rapport au contexte, la bonne syntaxe des schémas et des DTDs, ainsi que la cohérence des documents d'exemple au regard des modèles.

Pour cela, trois acteurs au moins se distinguent : l'utilisateur, le participant au groupe de travail et l'administrateur. L'utilisateur devra avoir accès aux modèles par le biais d'une interface Web à partir de n'importe quel navigateur. Un système de session personnel est mis en place. Cet utilisateur pourra être n'importe qui, mais il pourra être aussi un soumissionnaire d'un modèle dans la base. Le participant au groupe de travail est une personne qui se dote d'un mot de passe pour accéder à un espace de travail bien particulier. L'administrateur devra pouvoir gérer cette base de données (ajouter, modifier ou supprimer un champ), mais il devra aussi pouvoir modifier les différents modules du serveur.

Parmi les contraintes définies, on peut noter :

- aucun utilisateur et/ou propriétaire du modèle ne peut modifier un modèle qui ne lui appartient pas ;
- les propriétaires des modèles doivent être munis des mots de passe pour accéder à leurs espaces de travail ;
- les modèles déclarés doivent subir une opération de révision et de contrôle de la part d'un comité spécial pour assurer la cohérence des données ;
- l'application doit être ouverte et évolutive : facile à mettre à jour et à y ajouter d'autres fonctionnalités répondant à des nouveaux besoins ;
- ni le logiciel serveur, ni le logiciel client ne doit être d'une technologie propriétaire.

5-2 Structure des données

Avant de concevoir la structure de l'application, nous avons essayé de définir la structure des données qui permet de préciser tous les éléments, les entités, les attributs, les relations entre

eux, ainsi que les différentes caractéristiques de ceux-ci. Cependant, vu la nature de notre projet qui se veut à la fois générique touchant le maximum possible de domaines et à la fois spécifique à l'échange électronique des données d'affaires (commerce électronique), nous avons opté pour deux structures de données. La première répondant au premier objectif, c'est à dire une structure générale permettant à n'importe qui de l'utiliser pour déclarer son modèle de données, et la deuxième structure est spécifique à la déclaration des « core components³ ebXML » (composants élémentaires ebXML) [UN/CEFACT, 2001]. La définition de ces deux structures s'est fait en utilisant les schémas XML [Langlois, 2002].

La première structure se présente brièvement comme suit :

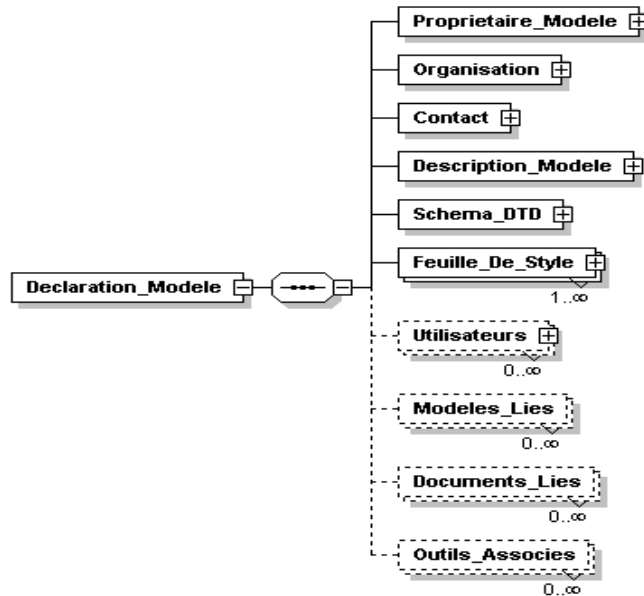


Figure 1 : Structure de données simple et générique

Quant à la deuxième structure, elle contient toutes les données de la première structure, mais elle englobe d'autres données liées à la définition d'un « core component ». Ces données sont extraites de la spécification ebXML [OASIS, 2002].

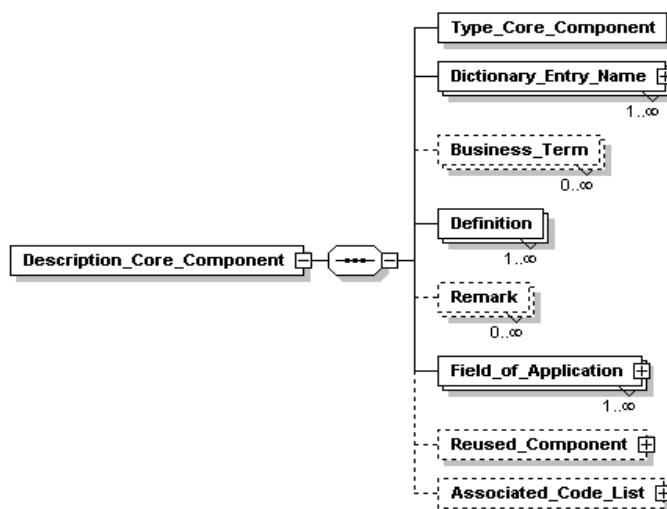


Figure 2 : Structure de données liée aux « core components ebXML »

³ Traduit en français *composants élémentaires*, les « core components » sont des pièces d'assemblage, ayant chacune une définition unique en sémantique d'affaire. Pour plus d'informations, veuillez voir : <http://www.autoroute.gouv.qc.ca/publica/normes/norme111.htm>

5-3 Structure de l'application

L'application est structurée en six parties :

- une partie contenant les éléments de définition du modèle ;
- une partie définissant toutes les Informations associées au modèle ;
- un module définissant tous les critères de recherche ;
- un module pour gérer le contrôle et la révision des modèles proposés à publication ;
- un module pour la consultation des données du répertoire ;
- et un module pour la soumission des modèles et documents XML.

Cette structure a été définie, après une étude de besoins qui a touché les partenaires principaux de ce projet à savoir Edifrance (Association pour le développement des échanges électroniques professionnels), MUTU-XML et GFII (Groupement Français de l'Industrie de l'Information), FING (Fondation Internet Nouvelle Génération), ainsi que l'UIC (Union Internationale de Chemins de Fer) à Bruxelles.

Cette structure peut être schématisée comme suit :

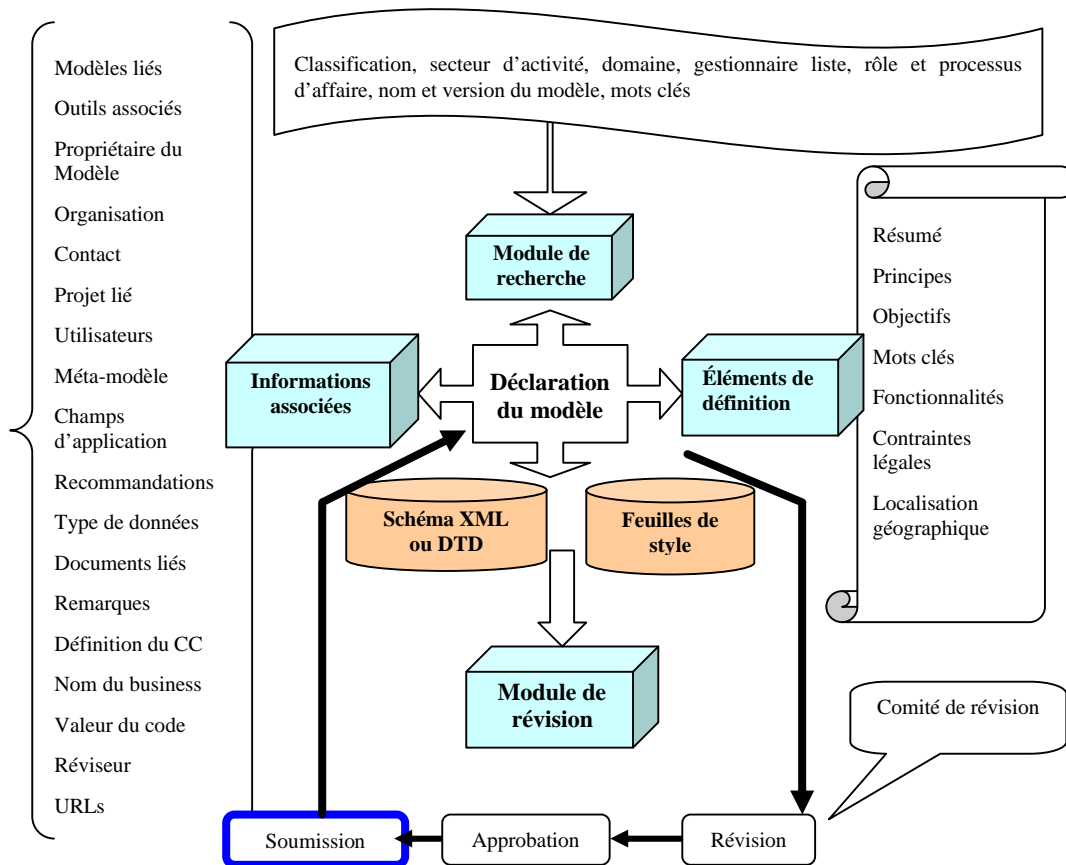


Figure 3 : Structure de la base de données du répertoire de modèles XML

Le module de recherche peut utiliser plusieurs critères : secteur d'activité, domaine d'usage, rôle dans le processus d'affaire, classification du produit, processus d'affaire, nom du modèle et mots clés. Ces critères peuvent être utilisés conjointement avec les opérateurs logiques "et" et "ou". Le module de recherche est accessible par consultation à distance. Les trois premiers critères constituent une liste fermée maintenue par le comité éditorial. Les autres sont ouverts.

Le module de consultation donne accès, à partir d'une recherche, aux données du répertoire. Il peut, soit permettre une consultation sur navigateur, soit permettre un téléchargement du résultat de recherche.

Quant au module de révision, il représente une antichambre pour assurer la meilleure diffusion des modèles XML et, donc, de favoriser la meilleure compréhension possible de l'objectif et des conditions d'utilisation d'un schéma particulier.

Le module de soumission d'une nouvelle entrée utilise le même modèle que celui qui sert à la restitution, à partir du module de recherche. La soumission est stockée en attente d'approbation. L'approbation permet de mettre à jour le répertoire.

En ce qui concerne la partie se rapportant aux éléments de définition du modèle, elle permet de définir le modèle à travers le résumé, les objectifs, les principes de base, les différentes fonctionnalités et les mots clés. Elle permet de définir aussi les caractéristiques des situations d'affaires influencées par des exigences légales ou réglementaires (lois, règlements, conventions, traités, ...).

L'autre partie permet, quant à elle, de présenter toutes les informations associées au modèle à savoir son propriétaire, l'organisation responsable, les modèles et documents liés, les outils utilisés, les principaux utilisateurs, etc.

5-4 Modes d'utilisation

L'utilisation de l'application « répertoire de modèles XML » diffère selon qu'on est soumissionnaire de modèle, utilisateur ou réviseur.

Pour un soumissionnaire, il doit passer par une identification (nom, prénom, Login et mot de passe, etc.) s'il est nouveau, si non, il n'entre que son login et son mot de passe pour accéder à une interface qui lui permet de choisir le type de modèle (modèle simple ou modèle des « core components ») qu'il souhaite enregistrer, ainsi que le mode de soumission (manuelle ou par exportation⁴), et il reçoit ainsi un formulaire qui lui permet de remplir toutes les données y compris le schéma XML. Ce système d'identification nous permet d'une part d'éviter les redondances des entrées au niveau propriétaires de modèles et d'autre part de contrôler et d'assurer que la modification d'un modèle ne se fait que par son propriétaire.

Pour un utilisateur du grand public qui veut consulter un modèle ou tous les modèles existants dans la base, il lance sa requête à partir de l'interface de recherche et reçoit une liste de modèles, avec seulement le nom du modèle et un lien vers son contenu global.

Quant au réviseur, il passe lui aussi par un mot de passe pour accéder aux modèles soumis, et après vérification de leurs contenus, il note ses commentaires et remarques sur chaque modèle pour son propriétaire si le modèle manque d'informations ou s'il ne répond pas aux exigences prédéfinies. Si non, il valide le modèle pour qu'il devienne accessible au grand public.

La première page de l'application nous permet donc de consulter la liste des modèles, de chercher un modèle par plusieurs critères ou de soumettre et/ou modifier un modèle en passant par une identification. Cette dernière utilisation (soumission et/modification) donne à l'utilisateur accès à une autre page sommaire qui lui propose quatre fonctionnalités (figure 4).

⁴ Lors de la soumission, on peut, au lieu de remplir toutes les données manuellement, les exporter à partir d'un fichier Excel, et ceci représente un besoin déclaré par les partenaires à l'état actuel qui travaillent beaucoup avec des tableurs Excel pour stocker leurs « core components ». Lors de la consultation et téléchargement, on peut faire l'opération inverse (importer les données de la base dans un fichier Excel).



Figure 4 : Page sommaire du répertoire de modèles XML

Comme le présente cette figure qui représente une des pages d'accès au répertoire de modèles XML, les principales fonctionnalités sont : la soumission, la consultation, la modification et l'exportation des modèles.

6 Résultats et perspectives

Le répertoire a été testé et est utilisé actuellement en local par les différents partenaires du projet notamment par Edifrance pour la publication des « core components ». Il sera bientôt hébergé et publié par l'UIC (Union Internationale des Chemins de fer) quand il aura évolué son serveur pour supporter les fonctionnalités des outils utilisés dans le développement de cette application. Ceci permettra aux différents partenaires du projet de publier leurs modèles et d'en profiter de l'existence d'autres modèles liés à différentes activités. Cette application permet également d'éviter certainement des travaux redondants d'une part, et encouragera d'autres à utiliser la technologie XML, tout en s'inspirant des modèles déjà publiés. Il offre, par ailleurs, un service d'aide à l'élaboration des schémas XML. L'évolution prévue de cette application consiste en premier temps à ajouter une fonctionnalité d'exporter les modèles de la base de données en format XML et de proposer après l'outil à différents acteurs utilisant XML tout en l'adaptant à leurs besoins.

7 Conclusion

Nous avons présenté dans cet article le répertoire de modèles de schémas XML. Le contexte de création et les structures de données et de l'application ont également été présentés, ainsi que quelques intérêts de cet outil. Cependant, d'autres intérêts peuvent être évoqués à travers cette application, à savoir le travail collaboratif, la recherche sémantique..., qui s'intègrent dans une problématique très large dépassant le cadre de cet article. Toutefois, nous pouvons dire que le répertoire de modèles XML représente un pas exceptionnel dans le domaine de l'EDI. Un tel projet était depuis un bon moment, (en tous cas, depuis le développement et l'introduction du métalangage XML dans le domaine du commerce électronique et dans

l'administration française) le souci de tous les organismes travaillant dans le domaine de la normalisation et la standardisation.

8 Références

- [Attar, 2002] ATTAR, Pierre. - Documentation Technique du modèle permettant l'échange de descriptions de modèles. - Version 0.1, 1^{er} Janvier 2002
- [Chauvet, 2002] CHAUVET, Jean-Marie. - Services Web avec SOAP, WSDL, UDDI, ebXML. - Paris : Eyrolles, 2002
- [GENCOD EAN, 2002a] GENCOD EAN & Edifrance. - Le commerce électronique pour l'entreprise : guide de mise en œuvre des échanges électroniques professionnels. - France, janvier 2002
- [GENCOD EAN, 2002b] GENCOD EAN & Edifrance.- comprendre XML pour les échanges électroniques professionnels (cd-rom). - France, mai 2002
- [Jospin, 2002] JOSPIN, Lionel. - Circulaire du 21 janvier 2002 relative à la mise en œuvre d'un cadre commun d'interopérabilité pour les échanges et la compatibilité des systèmes d'information des administrations, J.O n° 30 du 5 février 2002 page 2335, NOR: PRMX0205357C, Paris, 2002. - <http://www.legifrance.gouv.fr/WAspad/UnTexteDeJorf?numjo=PRMX0205357C>
- [Langlois, 2002] LANGLOIS, Marc ; MKADMI, Abderrazak. - Création d'un répertoire de modèles XML : cahier des charges, EDIFRANCE, 2002. - http://www.edifrance.org/accueil_schemaxml.php
- [OASIS, 2002] OASIS. - OASIS Registry/Repository Technical Specification, Working Draft 1.1 December 20, 2000.
- [Parent, 2002] PARENT, Richard. - Normes ouvertes en technologies de l'information : Spécification technique des composants élémentaires, Partie 1. - <http://www.autoroute.gouv.qc.ca/publica/normes/liste.htm>
- [UN/CEFACT, 2001] UN/CEFACT. - Core Components Technical Specifications, Part 1// Spécification technique des composants élémentaires, Partie 1, traduit par Richard Parent. - <http://www.autoroute.gouv.qc.ca/publica/normes/norme111.htm>
- [Van Der List, 2001] VAN DER VLIST, Eric.- Le répertoire de schémas XML de la MTIC ouvre ses portes. - <http://xmlfr.org/actualites/decid/010413-0002>

Sites Web

W3C : <http://www.w3.org/>

XML : <http://www.w3.org/XML/>

Schémas XML : <http://www.w3.org/XML/Schema>

ebXML : <http://www.ebxml.org/>.

OASIS : <http://www.oasis-open.org/>

Mutualiser l'effort de montée en compétences sur XML : <http://www.mutu-xml.org/>

EDIFRANCE : <http://www.edifrance.org>

GFII : <http://www.gfii.asso.fr/dep.htm>

FING : <http://www.fing.org/index.php?rubrique=lafing>

XML SPY : <http://www.xmlspy.com/>