



HAL
open science

Spécification déclarative pour des documents virtuels personnalisables

Sébastien Iksal, Serge Garlatti

► **To cite this version:**

Sébastien Iksal, Serge Garlatti. Spécification déclarative pour des documents virtuels personnalisables. Documents Virtuels Personnels 2002, 10-11 Juillet, Brest, ENST Bretagne, Jul 2002, Brest, France. pp.127 - 140. hal-02142950

HAL Id: hal-02142950

<https://hal.science/hal-02142950>

Submitted on 29 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SPECIFICATION DECLARATIVE POUR DES DOCUMENTS VIRTUELS PERSONNALISABLES

Sébastien Iksal, Serge Garlatti

Département d'Intelligence Artificielle et de Sciences Cognitives
ENST Bretagne
Technopôle de Brest Iroise – B.P. 832 – 29285 BREST Cedex
{prenom.nom}@enst-bretagne.fr

Résumé : Nous proposons un environnement de conception de documents virtuels personnalisables muni d'une spécification déclarative des processus de sélection et d'organisation intégrant les critères d'adaptation/personnalisation. La spécification du processus d'organisation est une description déclarative d'un document générique – un graphe orienté dont les arcs sont des relations sémantiques et les nœuds des composants sans contenu. La spécification de la sélection est réalisée en définissant les propriétés sémantiques du contenu associées à chaque nœud. Les critères d'adaptation sont spécifiés par des propriétés sémantiques associées aux documents génériques. Ces propriétés sont fondées sur des évaluations de fragments et des stéréotypes de lecteurs pour la gestion des stratégies d'adaptation. Le document générique est réutilisable, partageable et échangeable entre auteurs. Il est aussi aisément modifiable puisque sa spécification se compose de déclarations à un niveau sémantique. Il est ainsi possible de maintenir et de créer de nouveaux services en minimisant l'effort de conception et en maximisant la réutilisation – des fragments, des règles d'adaptation et des documents génériques.

Mots-clés : Composition, spécification, adaptation, ontologie, stéréotype

1 Introduction

De nombreux services sont accessibles de nos jours sur Internet comme par exemple, des portails, des cours en ligne, des musées virtuels, du e-commerce, des journaux en ligne, etc. Parmi les enjeux importants, on peut citer l'adaptation/personnalisation, la conception et l'évolution des documents ou services sur Internet, ainsi que le partage, la réutilisation, et l'échange d'informations. L'adaptation/personnalisation permet de prendre en compte des utilisateurs qui ont des besoins différents pour améliorer l'utilisabilité et la compréhension, et ont à faire face à un grand espace d'information. En fait, les services adaptatifs, encore appelés hypermédia adaptatifs – dans notre cadre -, peuvent fournir différents types d'information, différents mécanismes de navigation, différentes présentations et différentes organisations de services en fonction des besoins des utilisateurs [Brusilovsky 1996].

La notion d'hypermédia flexible et plus particulièrement celle de document virtuel permet de proposer des méthodes pour faciliter la conception, la création et la maintenance de services. Dès 1993, Thomas Gruber définit les documents virtuels comme des documents hypermédia qui sont générés à la demande en fonction de plusieurs sources d'information et en réponse à une demande de l'utilisateur (le lecteur) [Gruber, et al. 1996]. Pour Watters, un document virtuel est un document qui ne possède pas d'état persistant et pour lequel toutes les instances (documents réels) sont générées lors de la consultation [Watters, et al. 1999]. La génération de documents réels, appelée

composition, comporte au moins trois processus qui peuvent être séparés ou imbriqués : la sélection, l'organisation et l'assemblage des fragments. La spécification d'un document virtuel est constituée de l'ensemble des informations nécessaires au système pour composer le document réel. Le processus de composition peut être vu comme un mécanisme capable d'engendrer une quantité infinie de documents réels qui peut être utilisé pour différentes classes de services à partir du moment où les principes de la composition et de la spécification sont adéquats au service à concevoir. La réutilisation, le partage et l'échange d'informations et de documents au travers d'Internet nécessite des moteurs de recherche fondés sur le « Semantic Web » [Decker, et al. 1999, Decker, et al. 2000].

Nous avons étudié un environnement de conception de Documents Virtuels Personnalisables (DVP) qui possède une spécification déclarative des processus de sélection et d'organisation intégrant les critères d'adaptation/personnalisation. La spécification du processus d'organisation est une description déclarative d'un document générique – un graphe orienté dont les arcs sont des relations sémantiques et les nœuds des composants sans contenu. La spécification de la sélection est réalisée en définissant les propriétés sémantiques de contenu associées à chaque nœud. Les critères d'adaptation sont spécifiés par des propriétés sémantiques associées aux documents génériques. Ces propriétés sont fondées sur des évaluations de fragments et des stéréotypes de lecteurs pour la gestion des stratégies d'adaptation. Dans le cadre des projets ICCARS et CANDLE, un document générique représente le savoir-faire d'un auteur à un niveau sémantique. Il est fondé sur la théorie des structures rhétoriques ou RST [Mann, et al. 1988]. Ce document générique est réutilisable, partageable et échangeable entre auteurs. Il est aussi aisément modifiable puisque sa spécification se compose de déclarations à un niveau sémantique. Il est ainsi possible de maintenir et de créer de nouveaux services en minimisant l'effort de conception et en maximisant la réutilisation – des fragments, des règles d'adaptation et des documents génériques.

Dans une première partie, nous allons présenter la problématique des Documents Virtuels Personnalisables depuis la définition jusqu'à la spécification et la composition. Ensuite, nous proposerons une approche déclarative de la spécification et de la composition d'un DVP. En conclusion, nous reviendrons sur les apports de notre approche.

2 Problématique des Documents Virtuels Personnalisables

Les documents virtuels sont nés d'un besoin d'interactivité et d'individualisation des documents, plus particulièrement sur le Web. Si un utilisateur peut adapter/personnaliser le document réel produit en fonction de ses besoins, on parlera de document virtuel personnalisable, sinon uniquement de document virtuel. Un document virtuel peut être personnalisé à partir d'au moins un des trois critères suivants :

- Les contraintes utilisateurs sont des données non persistantes définies à partir d'une requête, d'un formulaire ou d'un simple hyperlien qui lui sont propres.
- Le modèle utilisateur : Il s'agit d'un ensemble de données rémanentes qui caractérisent un utilisateur ou un groupe d'utilisateurs particuliers. Un tel modèle peut contenir des caractéristiques sur les connaissances, les préférences, les objectifs, les centres d'intérêts, etc. d'un utilisateur [Brusilovsky 1996, Rich 1989].

-
- Le contexte [Brézillon 2002]: support de lecture, conditions géographiques, météorologiques [Garlatti, et al. 2000], etc.

En 1999, Crampes et Ranwez définissent les documents virtuels personnalisables comme les éléments et les mécanismes permettant de réaliser une construction dynamique de documents adaptatifs. Un DVP peut être considéré comme un ensemble d'éléments (ou fragments) associé à des mécanismes de filtrage, d'organisation et d'assemblage sous contraintes, i.e. en respectant un modèle de l'utilisateur et des principes narratifs [Ranwez, et al. 1999]. Bien souvent, l'organisation et la sélection de fragments pertinents nécessitent des mécanismes de recherche d'information et de filtrage qui utilisent entre autres les index des fragments. La caractérisation de ces fragments et leur indexation ne peuvent pas être séparées des mécanismes d'un DVP. En effet, un DVP est fondé sur deux étapes principales qui sont indissociables : la spécification et la composition du document réel. La spécification dépend bien évidemment des principes mis en œuvre dans la composition. Elle permet au minimum aux auteurs d'un DVP de définir les fragments (nature, granularité et indexation), mais aussi les règles d'organisation et d'assemblage. Les deux derniers ne sont pas toujours accessibles aux auteurs. Le choix de la granularité des fragments est important car elle intervient dans la composition du document, mais aussi dans l'indexation et la recherche de l'information. On peut distinguer deux niveaux de granularité l'un pour la recherche et l'autre pour la réutilisation [Vercoustre, et al. 1997]. En effet, un fragment peut être sélectionné, mais seule une partie de celui-ci peut être réutilisée. La composition a pour objet la construction d'un document réel à partir d'une spécification et des critères de personnalisation. Spécification et composition sont en quelque sorte deux faces d'une même pièce de monnaie ; elles ne peuvent être conçues l'une sans l'autre et forment un tout indissociable. Nous allons présenter l'étape de composition ainsi que les implications correspondantes en terme de spécification de document.

L'étape de composition varie d'un système à l'autre, néanmoins elle comporte au moins trois processus qui peuvent être séparés ou imbriqués : la sélection des fragments, l'organisation de ceux-ci et leur assemblage. Ils peuvent aussi être analysés tant du point de vue de la composition que du point de vue de la spécification.

2.1 Le processus de sélection

C'est un processus de recherche d'information qui doit apparier une spécification de contenu avec les index des fragments et les critères d'adaptation. La spécification du contenu est obtenue de deux façons :

- Soit par l'auteur qui construit des modèles de pages dont une partie est statique et dont l'autre est composée d'une ou plusieurs requêtes vers des bases d'information [Vercoustre, et al. 1997]. L'auteur peut aussi définir l'ensemble des objectifs à atteindre ([Ranwez 2000], projet Sibyl) et le système sélectionnera les fragments d'information permettant d'atteindre ces objectifs. Une autre possibilité consiste pour l'auteur à composer un schéma de nœuds, dont les nœuds comportent les informations nécessaires à l'identification de leur contenu (Requête dans une base de données, utilisation de méta-données) [Falquet, et al. 1999].
- Soit par le lecteur qui définit ses objectifs ([Ranwez 2000], projet Karina), une sorte de modèle de tâches à réaliser. Dans ce cas, le système recherche les fragments d'information permettant d'atteindre ces objectifs. Une autre

possibilité consiste à permettre au lecteur de définir son espace d'information et au système d'organiser cet espace par la suite (principe des moteurs de recherche).

Les caractéristiques des filtres peuvent, par exemple, être de type temporel - durée de consultation du fragment ou date de production du fragment [Ranwez 2000], de type sémantique – utilisation d'un modèle de domaine [Garlatti, et al. 1999, Nanard 1994], ou encore de type utilisateur – préférences du lecteur ou niveau de connaissance. Dans certains projets [Tazi, et al. 1999] [Balpe 1997], l'étape de sélection consiste à générer le contenu de façon dynamique.

2.2 Le processus d'organisation

Le processus d'organisation, appelé aussi ordonnancement [Ranwez 2000] ou planification [Tazi, et al. 1999], consiste à associer à l'espace d'information obtenu lors de l'étape de sélection, une ou plusieurs structures cohérentes et compréhensibles par le lecteur du document réel. L'organisation est le résultat de l'application de règles de composition qui peuvent, par exemple, prendre les formes suivantes:

- Une structure implicite résultat d'un calcul comme un tri – de type résultat de moteur de recherche, ordre temporel, etc. – Il peut s'agir de règles de précedence par l'utilisation de pré-requis, ou encore de règles de proximité sémantique – distance entre deux concepts dans un graphe - par la définition et l'utilisation, par exemple, d'un modèle de domaine [Ranwez 2000]. Il peut s'agir aussi de règles définissant les étapes à franchir pour atteindre un objectif comme les règles pédagogiques de Tazi [Tazi, et al. 1999].
- Une structure explicite comme un modèle de tâche, une structure narrative [Balpe 1997], etc. qui organise un site Web (portail, musée virtuel, etc.). Parmi les structures explicites, Falquet [Falquet, et al. 1999] propose l'utilisation d'un schéma de nœuds qui possèdent des expressions de contenu que le système utilise lors de l'étape de sélection de contenu. Vercoustre [Vercoustre, et al. 1997] propose une approche basée sur une structure logique de documents en SGML. Sa structure possède certains éléments statiques et d'autres qui correspondent à des spécifications utilisées pour rechercher l'information de façon dynamique (Requêtes dans des bases de données).

2.3 Le processus d'assemblage

Ce processus concerne toutes les opérations nécessaires afin de rendre le document plus facile à lire et à comprendre. Le système peut, par exemple, ajouter des transitions entre les fragments (appelées aussi liants [Ranwez 2000]) pour assurer la cohérence d'un ensemble de textes associés, ou encore effectuer des opérations sur les fragments (Fondu sur de la vidéo [Ranwez 2000]). Pour Rekik, l'auteur spécifie les opérations que le système peut effectuer sur les fragments (rotation pour une image, différents systèmes d'unités pour un graphique, etc.) [Rekik, et al. 1999]. Le processus d'assemblage concerne aussi la représentation de la structure globale du document ainsi que la représentation du document de façon visuelle.

3 Une approche déclarative

Nous proposons un environnement de conception de documents virtuels personnalisables qui possède une spécification déclarative des processus de sélection et d'organisation intégrant les critères d'adaptation/personnalisation. Il s'agit de concevoir une architecture ouverte de création de DVP. En d'autres termes, cet environnement doit être utilisable pour différents services sur Internet. Dans le cadre de nos activités, l'environnement est utilisé dans deux projets : ICCARS¹ et CANDLE². ICCARS a pour but d'étudier un système d'aide à la création de dossiers thématiques adaptatifs publiables sur Internet – tous les exemples seront pris dans ICCARS. Un dossier thématique est une synthèse faite par un ou plusieurs journalistes sur un sujet particulier. CANDLE a pour objet l'utilisation d'Internet pour permettre la collaboration entre les universités et l'industrie dans la création, le partage et la réutilisation de matériaux pédagogiques.

3.1 Hypothèses

Nos hypothèses principales sont les suivantes :

- Spécification déclarative des processus de sélection et d'organisation munis de critères d'adaptation à un niveau sémantique afin d'en faciliter la création, le partage, la réutilisation et l'échange : la spécification de l'organisation et de la sélection est une structure narrative définie par un auteur de façon à présenter un point de vue, un éclairage particulier d'un ensemble de fragments. Une structure narrative consiste en un graphe orienté composé de nœuds appelés composants qui contiennent une spécification de contenu et d'arcs qui sont des relations sémantiques typées (Théorie de la Structure Rhétorique ou RST [Mann, et al. 1988]) entre ces composants.
- Adaptation : comme le document réel fourni à l'utilisateur est un document hypermédia, nous nous sommes orientés sur les stratégies d'adaptation utilisées dans ce type de document. Elles sont classées en deux catégories [Brusilovsky 1996]: des stratégies de présentation adaptative et des stratégies de navigation adaptative. Nous proposons de gérer les cinq stratégies principales de navigation adaptative (annotation, guidage direct, masquage, masquage partiel et tri). Les stratégies de présentation adaptative sont déduites de celles-ci. Les stratégies de navigation adaptative ont pour objectif d'aider les utilisateurs à trouver leur chemin dans un vaste espace d'information et d'améliorer la compréhension en adaptant la présentation des liens ou en éliminant certains d'entre eux en fonction des besoins utilisateur. Des expérimentations ont montré qu'il n'est pas pertinent de fournir la stratégie d'annotation pour un novice, mais plutôt celle de masquage. L'annotation s'avère plus intéressante pour un expert [Specht, et al. 1999]. Il est donc nécessaire de définir des stéréotypes de lecteurs pertinents pour chaque stratégie.

¹ ICCARS : Integrated and Collaborative Computer Assisted Reporting System (<http://iccars.enst-bretagne.fr>)

² CANDLE : Collaborative And Network Distributed Learning Environment (<http://www.candle.eu.org>)

3.2 Principes généraux

Les processus de sélection, d'organisation et d'assemblage s'organisent de la façon suivante dans le moteur de composition :

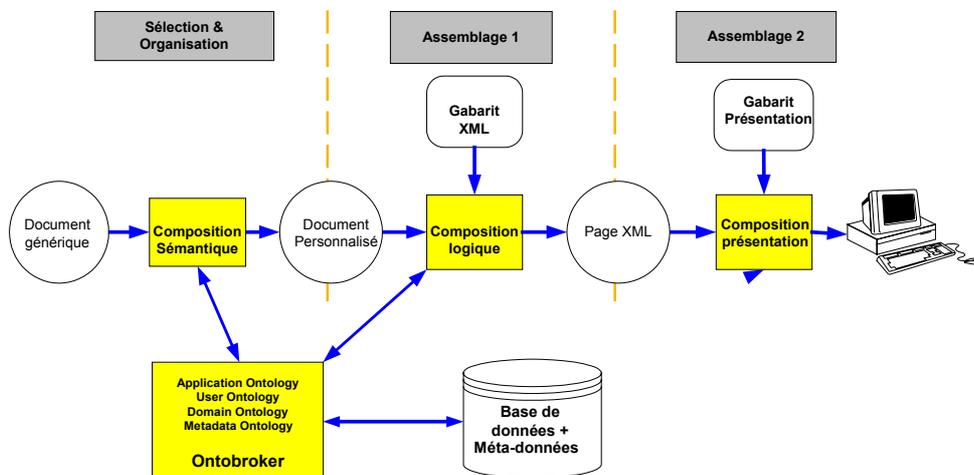


Figure 1: Architecture du moteur de composition

La composition sémantique a pour but de générer un document personnalisé à partir de l'espace d'information de ce document, de la stratégie de lecture choisie – structure narrative - et du modèle utilisateur. La structure narrative est un graphe, dont les nœuds contiennent une spécification du contenu extraite des méta-données associées aux fragments et dont les relations sémantiques sont typées et choisies par l'auteur. Le résultat de la composition sémantique est en fait, l'instanciation du document générique ou document personnalisé. Il s'agit donc d'un ensemble de fragments liés par des relations sémantiques. L'assemblage se décompose en deux étapes effectuées par la composition logique et la composition de présentation. La composition logique parcourt le document personnalisé et engendre pour chaque nœud (ou fragment) une page Web avec ses liens de navigation. Elle utilise un gabarit de page et le modèle utilisateur afin de générer une page XML. Cette page a pour contenu un fragment et les liens de navigation munis de propriétés pour gérer les stratégies d'adaptation [Brusilovsky, et al. 1996]. La composition de la présentation construit une page HTML en appliquant des règles de présentation dépendantes du lecteur et/ou de l'institution qui propose ces documents. Elle utilise les règles de l'auteur et/ou les préférences de l'utilisateur pour afficher les repères de navigation.

3.3 Les ontologies comme support à la déclarativité

Le moteur de composition génère un document réel à partir d'une structure narrative définie par un auteur et de l'espace d'information du document. La structure narrative est représentée par un modèle de document ainsi que par des types de fragments et de relations. Le modèle de document et le savoir-faire de l'auteur (types de fragments et de relations) sont regroupés dans une ontologie de l'application. La structure narrative est un graphe orienté dont les nœuds contiennent des spécifications permettant de retrouver les fragments d'informations. Ces fragments sont indexés par des méta-données sémantiques dont certaines prennent leur valeur dans l'ontologie de domaine et dans celle de l'application. La personnalisation du contenu se fait par filtrage à partir d'un modèle de

l'utilisateur. Le modèle de l'utilisateur nous renseigne sur les préférences et les connaissances de ce dernier exprimées en fonction du modèle de domaine.

L'architecture de composition est fondée sur quatre ontologies, une ontologie de l'application, un schéma de méta-données, un modèle du domaine ainsi qu'une modélisation de l'utilisateur. La gestion des ontologies au sein du mécanisme de composition sémantique repose sur OntoBroker [Fensel, et al. 1999] qui possède un système de requête fondé sur un moteur d'inférence. OntoBroker est un logiciel de gestion de connaissances qui permet le filtrage et la recherche d'information dans l'espace d'information – les fragments - comme dans les ontologies.

4 Spécification et Composition de DVP

Dans notre contexte, la spécification d'un DVP est constituée de celles du document générique, de l'adaptation et de l'assemblage. Les étapes de sélection et d'organisation sont fortement liées puisque la sélection prend place dans un nœud d'une structure narrative afin d'en assurer la cohérence. Nous allons d'abord présenter les spécifications du document générique, de l'adaptation et enfin de l'assemblage, puis ensuite les différentes étapes de composition.

4.1 La spécification du document générique

Un document est composé d'un ensemble de fragments et d'une ou plusieurs organisations de ceux-ci (Fig. 2) – structures narratives. Les fragments peuvent être de deux types : abstrait ou atomique. Un fragment abstrait est composé d'un ensemble de fragments abstraits ou atomiques et d'une ou plusieurs organisations. Un fragment atomique est une unité d'information non décomposable.

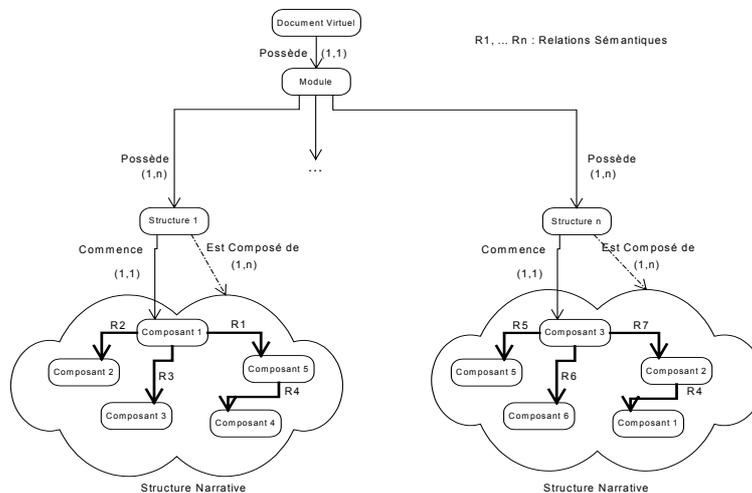


Figure 2: Modèle de document

La modélisation du document est définie comme suit : le document générique se compose d'un ensemble de fragments sélectionnés par l'auteur – associés explicitement afin de définir l'espace d'information du document -, et d'une ou plusieurs structures narratives entre ces fragments. Une structure narrative est un graphe orienté sans circuit représenté par un ensemble de composants et de relations. L'un des composants a le statut particulier de racine de la structure. Les composants sont reliés entre eux par une relation sémantique provenant de la RST. Pour chacun des composants, l'auteur indique les

fragments pertinents. La spécification du contenu est générée par le système auteur qui ne retient qu'un sous-ensemble des méta-données permettant de retrouver ces fragments et uniquement ceux-là. D'autres méta-données sont utilisées pour déterminer à la composition les fragments variants et serviront donc pour les stratégies d'adaptation.

Le schéma de méta-données est composé de six parties : **Générale**: Informations d'identification de la ressource : le titre, les auteurs, etc. ; **Cycle de vie**: Informations relatives à la gestion des différentes versions d'une ressource, comme le numéro, la date, les auteurs, etc. ; **Méta Méta-données**: Informations sur les méta-données elles-mêmes, qui les a créées, validées, langue et date ; **Droits**: Informations relatives aux conditions d'utilisation de la ressource (coût, licence, restrictions d'accès, etc.) ; **Technique**: Informations d'ordre technique sur la ressource comme son format, sa taille, etc. ; **Classification**: Description du contenu : à chaque concept du domaine, niveau de connaissance (Novice, Intermédiaire, Expert) requis pour comprendre la ressource.

Les méta-données utilisées afin de spécifier le contenu des nœuds correspondent au MD.6.1.1 et MD.6.2.1 de la table 1, c'est-à-dire les concepts provenant du domaine qui identifient le contenu du fragment et les concepts de l'application identifiant les types de fragments – Interview, compte-rendu, etc. en ce qui concerne l'application journalistique ICCARS.

MD.6	Classification		Domaine de valeur
MD.6.1	Domaine	Liste	
MD.6.1.1	Concept	Valeur unique	Ontologie de domaine
MD.6.1.2	Niveau	Valeur unique	Ontologie des utilisateurs : Niveaux
MD.6.2	Application	Valeur unique	
MD.6.2.1	Type de fragment	Valeur unique	Ontologie de l'application : types de fragments

Table 1. Extrait des méta-données

Dans le cas où l'indexation du contenu ne suffit pas à spécifier le contenu du nœud, lorsque par exemple plusieurs fragments non prévus par l'auteur s'apparient avec la spécification, d'autres méta-données peuvent être prises en compte dans les sections : générale, cycle de vie, et technique.

4.2 La spécification de l'adaptation

4.2.1 Principes

Le système est conçu pour prendre en compte cinq stratégies de navigation adaptative et la présentation adaptative qui en découle [Brusilovsky 1996, De Bra, et al. 1998]. Une stratégie de navigation adaptative classe et filtre les liens du document hypermédia. Les cinq stratégies de navigation adaptative mises en oeuvre sont les suivantes :

- Annotation : chaque lien est présenté différemment en fonction du niveau de pertinence qui lui est associé. Une image ou un code de couleur peut être associé à un lien à partir d'une évaluation du lien qui prend en compte les caractéristiques du lecteur. On pourra, par exemple, mettre un feu vert devant les liens les plus recommandés pour le lecteur, un feu orange pour ceux qu'il peut consulter mais avec lesquels il aura des difficultés et un feu rouge pour ceux qui ne lui correspondent pas. Les liens sont donc classés.
- Guidage direct : le système détermine le « meilleur lien suivant » à partir du fragment courant. Les liens sont donc filtrés.

- Tri : les liens sont triés en fonction de leur niveau de pertinence pour un utilisateur particulier. Le système placera le lien le plus pertinent au sommet de la liste. Les liens sont donc classés.
- Masquage : les liens estimés les plus pertinents sont retenus pour l'utilisateur. Les liens sont donc filtrés.
- Masquage partiel : le système affiche les liens qui possèdent un niveau de pertinence compris dans un intervalle choisi par l'auteur. Les liens sont donc filtrés et classés. Les liens retenus sont ensuite annotés.

Pour toutes les stratégies d'adaptation, il est nécessaire d'évaluer la pertinence de chaque lien ou du fragment pointé par ce lien afin de le classer et/ou de le filtrer. Généralement, les hypermédia adaptatifs proposent un maximum de 5 classes, au delà, le lecteur éprouve des difficultés à se repérer. Nous avons choisi d'évaluer les fragments pointés par un lien. De plus, un auteur peut choisir de 2 à 5 classes, disjointes et totalement ordonnées, de liens pour un document générique. Ces cinq classes sont nommées : very good, good, average, bad, very bad. Pour chaque classe choisie par l'auteur, les propriétés nécessaires et suffisantes d'appartenance sont spécifiées.

4.2.2 L'évaluation des fragments

Les règles d'appartenance à une classe sont définies par des expressions booléennes qui peuvent prendre en compte un stéréotype de lecteur (des propriétés du modèle utilisateur, ontologie de l'utilisateur), un sous-ensemble des méta-données dédié aux fragments variants – ontologie des méta-données - et un niveau de connaissance des concepts du domaine – ontologie du domaine – calculé à partir des méta-données et du modèle utilisateur (Pourcentage de concepts connus, pourcentage de concepts de niveau suffisant). La grammaire (Table 2.) détaille le mode d'écriture de la règle d'appartenance par l'auteur. Cette règle est traduite automatiquement par le système dans le langage de la F-Logic.

```

Regle ::= Condition_Booleenne |
        Condition_Booleenne Condition_Supplementaire+;
Condition_Booleenne ::= (Objet Comparateur Valeur) |
                        Operateur_Unaire (Objet Comparateur Valeur);
Condition_Supplementaire ::= Operateur_Binaire Condition_Booleenne;
Operateur_Unaire ::= NOT;
Operateur_Binaire ::= AND | OR;
Objet ::= User Model Features | Metadata entries | PCTAGEKNOWN |
        PCTAGEGOODLEVEL | ...;
Comparateur ::= = | < | > | <= | >= | <> | IN;
Valeur ::= STRING | NUMBER;

```

Table 2. Grammaire de définition de règles d'évaluation

La règle est une condition qui doit être satisfaite par le fragment et le lecteur pour que l'état correspondant soit associé au fragment. Un exemple de règle et sa signification :

```
Good = "(JOB = Fisherman) AND (PCTAGEKNOWN >= 75) "
```

Le fragment est évalué à Bon si le lecteur est un pêcheur et qu'il connaît au moins 75% des concepts du domaine qui indexent le fragment.

Pour le niveau de connaissance, nous distinguons deux cas possibles : le concept connu et le concept ayant au moins le niveau requis. En effet, lors de l'indexation il est demandé à l'auteur de définir un niveau minimal requis pour la compréhension du fragment (ou

document) sur chacun des concepts du domaine qui sert à l'indexation. Cette même démarche est réalisée lors de la création et lors de l'évolution du modèle du lecteur. Les éléments calculés sont :

- PCTAGEKNOWN : Pourcentage de concepts connus
- PCTAGEGOODLEVEL : Pourcentage de concepts dont la connaissance est évaluée (ou estimée) au moins au niveau requis lors de l'indexation – méta-données –

Pour chaque document générique, les propriétés sémantiques suivantes sont définies pour l'évaluation des liens : nombre de classes, règles d'appartenance pour chaque classe. Pour la gestion de stratégies d'adaptation, des propriétés spécifiant pour chaque stratégie d'adaptation le stéréotype de lecteur pertinent et le mode de gestion des classes de liens sont également associées à chaque document générique. L'objet du mode de gestion est de déterminer les classes de liens conservés et/ou supprimées pour chaque stratégie. Pour les stratégies, guidage directe, annotation et tri, le mode de gestion est toujours le même et en quelque sorte implicite: guidage direct, on conserve la meilleure classe, les autres sont supprimées ; annotation, toutes les classes sont conservées ; le tri, toutes les classes sont conservées et l'ordonnement se fait de la meilleure à la moins bonne. Par contre, pour le masquage et le masquage partiel, il est nécessaire de définir quelles sont les classes conservées. Toutes ces propriétés associées à un document générique ainsi que celles des méthodes d'adaptation sont déclarées dans le modèle de document – ontologie de l'application.

4.2.3 Stéréotype de lecteur et stratégies d'adaptation

Un lecteur bénéficie d'une stratégie d'adaptation si son modèle et le stéréotype associé à la stratégie s'apparient. Pour la définition des stéréotypes de lecteurs, l'auteur peut utiliser des caractéristiques provenant de l'ontologie de l'utilisateur comme, par exemple dans ICCARS, l'âge ou l'activité professionnelle, ou encore des caractéristiques plus fines comme la connaissance ou le niveau de connaissance de tel ou tel concept de l'ontologie du domaine ;

```

Stereotype ::= Condition_Booleenne |
              Condition_Booleenne Condition_Supplementaire+ ;
Condition_Booleenne ::= (Objet Comparateur Valeur) |
                        Operateur_Unaire (Objet Comparateur Valeur);
Condition_Supplementaire ::= Operateur_Binaire Condition_Booleenne;
Operateur_Unaire ::= NOT ;
Operateur_Binaire ::= AND | OR ;
Objet ::= Age, Location, Job, Domain Concept/Level, ...;
Comparateur ::= = | < | > | <= | >= | <> | IN ;
Valeur ::= STRING | NUMBER | Niveau;

```

Table 3. Grammaire de définition de stéréotype

La grammaire (Table 3.) détaille le mode d'écriture du stéréotype par l'auteur. Ce stéréotype est traduit automatiquement par le système dans le langage de la F-Logic [Ontoprise 2002].

Un exemple de stéréotype et sa signification :

```

ANNOTATION = "( (AGE < 18) OR (JOB = Student)) AND
              (LOCATION IN Bretagne)"

```

Les différents "objets" utilisés pour l'élaboration du stéréotype sont des caractéristiques du modèle de lecteur. Pour les propriétés sur la connaissance, il s'agit d'un modèle de recouvrement en référence à l'ontologie du domaine. Pour cette caractéristique, la condition porte sur le niveau de connaissance que possède le lecteur sur un concept de l'ontologie du domaine (par exemple, Novice sur la peinture antifouling).

La présentation adaptative est obtenue par l'utilisation de stratégies de navigation adaptative. En effet, la navigation adaptative agit sur les liens. Etant donné que la granularité de recherche est identique à celle de la réutilisation, un fragment ne peut être que conservé ou supprimé. Lorsqu'il y a filtrage des liens – masquage, guidage direct et masquage partiel -, les fragments correspondants aux liens supprimés ne sont pas conservés. Il y a donc présentation adaptative (ou adaptation du contenu) et modification de la structure narrative.

4.2.4 L'assemblage

L'assemblage en ce qui nous concerne correspond à la mise en place de repères de navigation locaux et globaux [Garlatti, et al. 1999] ainsi qu'à la mise en place de stratégies d'adaptation de la navigation et de la présentation. La gestion de repères de navigation se fait au travers de gabarits de pages, ces gabarits sont déclarés soit par l'auteur, soit par l'organisme pour lequel travaille l'auteur et sont instanciés dans l'étape de composition. Il s'agit de pages XML dans lesquelles des composants correspondants à différents services (repères de navigation, service de recherche d'information, etc.) sont organisés. Même si le lecteur peut définir ses préférences en terme de stratégies d'adaptation au sein de son modèle, ce sont les contraintes de l'auteur qui sont prioritaires sur les préférences des lecteurs. Le système utilisera les préférences du lecteur pour choisir celle qui sera effectivement appliquée.

4.3 Composition des Documents Réels

Après une étape de recherche de documents et un choix de structure narrative pour un document générique donné, le système recherche les fragments pertinents, les organise et finalement les assemble de façon à obtenir, soit une page, soit un site Web (un ensemble de pages reliées entre elles).

4.3.1 L'organisation et la sélection du contenu

La composition sémantique parcourt la structure narrative sélectionnée par le lecteur, instancie les composants soit intégralement, soit au fur et à mesure de l'avancement du lecteur dans le parcours – les deux options sont offertes. L'instanciation consiste à utiliser les spécifications de contenu et le schéma de méta-données afin de sélectionner les fragments concernés dans l'espace d'information du document. Le système recherche dans l'espace d'information l'ensemble des fragments qui sont indexés par tous les concepts – ou leurs sous-concepts - inclus dans la spécification des composants à l'aide des entrées MD.6.1.1 et MD.6.2.1 de la table 1, ou d'autres éventuellement. Cette sélection est obtenue par l'intermédiaire de requêtes au moteur d'inférence d'OntoBroker. Le résultat de l'étape de sélection est un ensemble de fragments variants. Ils diffèrent soit par le niveau de connaissance requis pour la compréhension, soit par d'autres critères plus techniques comme le type de fragment (image, texte, vidéo, bande sonore, etc.), ou la taille. L'instanciation d'une structure narrative passe aussi par l'ajout et la suppression de

liens – donc de fragments en fonction des stratégies d'adaptation autorisés pour le lecteur courant.

La gestion des liens se fait de la façon suivante :

- Soit A et B deux nœuds et R1 un lien entre A et B. Lors de l'instanciation du nœud B, si plusieurs fragments sont obtenus, le système crée autant de liens R1 que de fragments dans B. Il s'agit de plusieurs liens de même type et de même nœud source (A). Le système évalue chaque fragment et le classe – association d'un état unique. Cet état est utilisé ensuite par le processus de navigation adaptative.

En ce qui concerne l'évaluation des fragments, le système utilise les règles d'évaluation définies par l'auteur ainsi que le modèle du lecteur. Dans un premier temps, le système calcule systématiquement - pour des raisons techniques dues à OntoBroker - pour chaque fragment deux valeurs : Le pourcentage de concepts connus par le lecteur et le pourcentage de concepts pour lesquels le lecteur possède le niveau de connaissance requis par l'auteur. Ensuite, le système vérifie si le modèle du lecteur et le fragment correspondent à l'un des stéréotypes définis par l'auteur, si tel est le cas, l'état de pertinence correspondant au stéréotype est associé au fragment.

4.3.2 L'assemblage

L'assemblage se décompose en deux étapes successives : l'une effectuée par le moteur de composition logique qui détermine le contenu de chaque page Web et l'autre par le moteur de composition de la présentation qui gère l'aspect visuel de la page. En fonction de la politique d'adaptation choisie par l'auteur pour le lecteur courant et/ou les préférences du lecteur, la page XML engendrée par le moteur de composition logique contiendra plus ou moins de liens vers d'autres fragments. Aux liens conservés, une propriété est associée afin que le moteur de composition de la présentation puisse effectuer l'affichage requis de ces liens. Le moteur de composition de la présentation calculera une présentation de page HTML à partir d'un gabarit - qui correspond au gabarit de page XML utilisé à l'étape précédente – et des propriétés associées aux liens de navigation.

4.3.3 Mise en œuvre de l'environnement de conception de DVP

Même si l'utilisation d'ontologies comme « paramètres » du moteur de composition permet d'engendrer des documents réels pour de nombreux services différents, ce n'est pas suffisant pour assurer l'utilisation de tels méthodes de développement de services. Il est nécessaire d'y ajouter des outils de création auteur pour la spécification qui permettent de s'affranchir des langages formels utilisés et de la complexité des processus mis en œuvre. Il faut que ces outils soient les plus transparents possibles pour l'auteur. C'est pourquoi nous avons développé des outils associés à cette approche tels que l'outil de création de structures narratives, l'outil d'indexation basé sur le schéma de méta-données, etc. Presque tous les outils développés possèdent des interfaces engendrés dynamiquement à partir des ontologies disponibles. Ils sont donc réutilisables pour d'autres services à partir du moment où la structure globale du service est un graphe orienté. En effet, il suffit de changer les ontologies, et les gabarits de construction logique et de transformation de présentation.

5 Conclusion et perspectives

Nous avons proposé un environnement déclaratif pour la spécification et la composition d'un DVP qui intègre plusieurs stratégies d'adaptation. Il est basé sur l'utilisation de quatre modèles que nous avons représentés sous forme d'ontologies : Un modèle de document, un modèle de domaine, un modèle utilisateur ainsi qu'un schéma de méta-données. Le partage et la réutilisation des fragments est assuré par le schéma de méta-données ainsi que le modèle de domaine qui sont eux-même partageables et accessibles sur Internet. Grâce à l'utilisation de standards tels que ceux définis par le « Semantic Web », les fragments peuvent donc être réutilisés et partagés avec d'autres.

Le passage d'un service à d'autres passe par la définition de nouveaux modèles. Par exemple, le passage des dossiers thématiques sur le domaine de la pêche (ICCARS) aux cours sur les réseaux (CANDLE) a nécessité uniquement la réécriture du modèle de domaine, la modification du modèle utilisateur et l'ajout d'une section pédagogique dans les méta-données. En ce qui concerne les stratégies d'adaptation, les états associés aux fragments ont été changés. Mais, comme l'adaptation est également gérée de façon déclarative, c'est l'auteur qui définit les cadres d'application de ces stratégies. En effet, lors de la création d'un document, l'auteur associe à chaque stratégie un stéréotype de lecteur concerné par cette stratégie. Ce stéréotype dépend du modèle utilisateur qui lui aussi est déclaré. Il est tout à fait envisageable d'accroître le nombre des méthodes d'adaptation en modifiant une partie de l'ontologie correspondante.

Le principe d'évaluation des fragments et de gestion des stratégies d'adaptation est en cours d'extension pour offrir une tâche de recherche d'information adaptative – fragments et documents génériques - fondée sur le modèle utilisateur – lecteur ou auteur – et les méta-données des fragments. Il s'agit par exemple d'utiliser les caractéristiques « centres d'intérêts » et « connaissances » d'un utilisateur afin d'appliquer des stratégies d'adaptation combinées utilisant chacune des caractéristiques différentes comme par exemple : tri en cinq classes de liens à partir des « centres d'intérêts » plus annotation des liens de chaque classe sur les « connaissances » en 3 sous-classes.

6 Références

- [J. Balpe 1997] J. Balpe, "Trois mythologies et un poète aveugle", *Colloque "De la différence des arts"*, Paris, IRCAM, p.
- [P. Brézillon 2002] P. Brézillon, "Hors du contexte, point de salut", *Séminaire "Objets Communicants"*, Autrans, France, p.
- [P. Brusilovsky 1996] P. Brusilovsky, "Methods and techniques of adaptive hypermedia", *User Modeling and User-Adapted Interaction*, vol. 6, n° 2-3, 1996, p. 87-129,
- [P. Brusilovsky, E. Schwarz and G. Weber 1996] P. Brusilovsky, E. Schwarz and G. Weber, "ELM-ART: An intelligent tutoring system on World Wide Web", *Third International Conference on Intelligent Tutoring Systems, ITS-96*, Montreal, p. 261-269.
- [P. De Bra and L. Calvi 1998] P. De Bra and L. Calvi, "AHA! An open Adaptive Hypermedia Architecture", *The New Review of Hypermedia and Multimedia*, vol. 4, n° 1998, p. 115-139,
- [S. Decker, M. Erdmann, D. Fensel and R. Studer 1999] S. Decker, M. Erdmann, D. Fensel and R. Studer, "Ontobroker: Ontology Based Access to Distributed and Semi-Structured Information", *Conference on Database Semantics*, Rotorua, New Zealand, Kluwer Academic Publishers, p.
- [S. Decker, D. Fensel, F. V. Harmelen, I. Horrocks, S. Melnik, M. Klein and J. Broekstra 2000] S. Decker, D. Fensel, F. V. Harmelen, I. Horrocks, S. Melnik, M. Klein and J. Broekstra, "Knowledge Representation on the Web", On to Knowledge Project.

- [G. Falquet, L. Nerima, J. Guyot, C. Vanoirbeek and Y. A. Rekik 1999] G. Falquet, L. Nerima, J. Guyot, C. Vanoirbeek and Y. A. Rekik, "Des documents virtuels pour lire les bases de données", *Workshop sur les Documents Virtuels Personnalises, IHM'99*, Montpellier, 23 Novembre, p.
- [D. Fensel, A. Witt, J. Angele, S. Decker, M. Erdmann, H. P. Schnurr, S. Staab and R. Studer 1999] D. Fensel, A. Witt, J. Angele, S. Decker, M. Erdmann, H. P. Schnurr, S. Staab and R. Studer, "On2broker in a Nutshell", *the 8th World Wide Web Conference*, Toronto, p. May 11-14.
- [S. Garlatti and S. Iksal 1999] S. Garlatti and S. Iksal, "Documents virtuels personnalisables pour des systèmes d'informations en ligne", *Workshop sur les Documents Virtuels Personnalises, IHM'99*, Montpellier, 23 Novembre, p. 22-26; <http://www.site-eeerie.ema.fr/~multimedia/ihm99/>.
- [S. Garlatti and S. Iksal 2000] S. Garlatti and S. Iksal, "Concept Filtering and Spatial Filtering in an Adaptive Information System", Trento, Springer Verlag, p. 315-318, 2000.
- [S. Garlatti, S. Iksal and P. Kervella 1999] S. Garlatti, S. Iksal and P. Kervella, "Adaptive On-line Information System by means of a Task Model and Spatial Views", *Second Workshop on Adaptive Systems and user Modeling on the World Wide Web*, Toronto and Banff, Eindhoven University of Technology, p. 59-66.
- [T. R. Gruber, S. Vemuri and J. Rice 1996] T. R. Gruber, S. Vemuri and J. Rice, "Model-based Virtual Document Generation", May, Knowledge Systems Laboratory.
- [W. C. Mann and S. A. Thompson 1988] W. C. Mann and S. A. Thompson, "Rhetorical Structure Theory: Toward a functional theory of text organization", *Text*, vol. 8, n° 3, 1988, p. 243-281,
- [J. Nanard 1994] J. Nanard, "Multimédia and Hypermédia facing the concept of active document", *BIWIT'94*, Bayonne, p. 231-243.
- [Ontoprise 2002] Ontoprise, "How to Write F-Logic Programs - A Tutorial for the Language F-Logic", Tutorial, Mai,
- [S. Ranwez 2000] S. Ranwez, "Composition Automatique de Documents Hypermédia Adaptatifs à partir d'Ontologies et de Requêtes Intentionnelles de l'Utilisateur", PhD, Université de Montpellier 2, 2000.
- [S. Ranwez and M. Crampes 1999] S. Ranwez and M. Crampes, "Conceptual document and hypertext documents are two different forms of virtual document", *Workshop on Virtual Document, Hypertext Functionality and the Web*, Toronto, p.
- [Y. A. Rekik, C. Vanoirbeek, G. Falquet and L. Nerima 1999] Y. A. Rekik, C. Vanoirbeek, G. Falquet and L. Nerima, "Reusing Dynamic Document Fragment through Virtual Documents: Key Issues In Document Engineering", *Workshop sur les Documents Virtuels Personnalises, IHM'99*, Montpellier, 23 Novembre, p. 32-36.
- [E. Rich 1989] E. Rich, "Stereotypes and user modeling", berlin, Springer verlag, p. 35-51, 1989.
- [M. Specht and A. Kobsa 1999] M. Specht and A. Kobsa, "Interaction of domain expertise and interface design in adaptive educational hypermedia", *Second Workshop on Adaptive Systems and User Modeling on the World Wide Web*, Toronto & Banff, Eindhoven University of Technology, p. 89-93.
- [S. Tazi and Y. Altawki 1999] S. Tazi and Y. Altawki, "Création de documents virtuels - Cas des supports de cours", *Workshop sur les Documents Virtuels Personnalises, IHM'99*, Montpellier, 23 Novembre, p.
- [A.-M. Vercoustre, J. Dell'oro and B. Hills 1997] A.-M. Vercoustre, J. Dell'oro and B. Hills, "Reuse of Information through virtual documents", *Second Australian Document Computing Symposium*, Melbourne, Australia, p.
- [C. Watters and M. Shepherd 1999] C. Watters and M. Shepherd, "Research issues for virtual documents", *Workshop on Virtual Document, Hypertext Functionality and the Web*, Toronto, p.