HAL
open science

Adaptation et personnalisation dans le web sémantique

Yannick Prie, Serge Garlatti

To cite this version:

Yannick Prie, Serge Garlatti. Adaptation et personnalisation dans le web sémantique. Rapport de l'action spécifique CNRS/STIC sur le Web sémantique. (J. Charlet, P. Laublet, C. Reynaud, Eds), pp.79 - 89, 2003. hal-02142948

HAL Id: hal-02142948

https://hal.science/hal-02142948v1

Submitted on 3 Jul 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.

6 – Adaptation et personnalisation dans le Web sémantique

S. Garlatti[1], Y. Prié[2]

[1]Département IASC GET - ENST Bretagne
Serge.Garlatti@enst-bretagne.fr

[2] LIRIS FRE 2672 CNRS
Université Claude Bernard Lyon 1
yprie@liris.univ-lyon1.fr

Résumé

Au travers d'Internet, un nombre potentiellement infini de services et de documents est accessible à tous les usagers. La plupart des services et documents fournis actuellement sur Internet proposent une organisation, un contenu, un mode d'interaction et une présentation uniques pour tous. Ceci peut être suffisant dans certains cas. Mais tous les utilisateurs ne sont pas intéressés par les mêmes informations et n'ont pas les mêmes attentes, connaissances, compétences, centres d'intérêts, etc. Ils ne sont capables de comprendre ou d'accepter que des services et des documents dont l'organisation, le contenu, les modes d'interaction et la présentation sont adaptés à leurs besoins. Parmi les problèmes qui se posent on peut citer : l'accès à l'information pertinente, la navigation dans un grand espace de ressources et la compréhension d'une ressource complexe. L'adaptation / personnalisation dans le «Web sémantique » peut s'appuyer sur les domaines suivants : les hypermédia adaptatifs, la modélisation utilisateur et les documents virtuels personnalisables. Les travaux actuels sur l'adaptation dans le Web sémantique peuvent être décomposés en deux parties : l'une sur les hypermédias adaptatifs et l'autre sur les documents virtuels adaptatifs. Les futures recherches sur le Web Sémantique pour l'adaptation / personnalisation peuvent s'articuler autour des cinq points suivants : la modélisation utilisateur et des utilisations, la protection de la vie privée, la recherche d'information, les services et documents adaptatifs, et l'ergonomie des services et des documents.

1 PRÉSENTATION ET IMPORTANCE DE LA PROBLÉMATIQUE DU POINT DE VUE DES USAGES

1.1 ADAPTATION/PERSONNALISATION SUR LE WEB SÉMANTIQUE

Au travers d'Internet, un nombre potentiellement infini de services et de documents est accessible à tous les usagers. La plupart des services et documents fournis actuellement sur Internet propose une organisation, un contenu, un mode d'interaction et une présentation uniques pour tous. Ceci peut être suffisant dans certains cas. Mais tous les utilisateurs ne sont pas intéressés par les mêmes informations et n'ont pas les mêmes attentes, connaissances, compétences, centres d'intérêts, etc. Ils ne sont capables de comprendre ou d'accepter que des services et des documents dont l'organisation, le contenu, les modes d'interaction et la présentation sont adaptés à leurs besoins. Parmi les problèmes qui se posent on peut citer :

- L'accès à l'information pertinente : la recherche d'information sur Internet par des mots clés n'est pas satisfaisante. En effet, les critères de rappel et de précision ne sont pas satisfaisants (précision souvent inférieure à 1%). L'utilisation d'ontologies et donc du « web sémantique » va grandement améliorer les résultats. Néanmoins, il restera bien souvent un nombre considérable de ressources qui satisferont les demandes des utilisateurs. Il est donc nécessaire de mieux cibler les résultats en fonction des besoins utilisateurs afin d'effectuer un filtrage des ressources.
- La navigation dans un grand espace de ressources : dès que l'espace d'information accessible à un utilisateur par navigation est important, la perte de repères et la désorientation ne permettent pas à l'utilisateur de retrouver les informations recherchées ni de savoir d'où il vient, où il peut aller et comment y aller. Il est important ici de pouvoir guider l'utilisateur dans cet espace d'information en d'autres termes d'adapter l'accès à l'information en fonction de ses besoins.
- La compréhension d'une ressource complexe – site web, service, etc. : en psychologie cognitive, la compréhension d'un document est souvent caractérisée par la construction mentale d'une représentation, d'un modèle de ce document. La lisibilité du document peut être définie comme l'effort mental nécessaire au processus de construction d'un modèle (Kaheneman, 1973; Thüring, Hannemann *et al.*, 1995). Afin d'accroître la lisibilité d'un document hypermédia, il est nécessaire d'assister l'utilisateur dans la construction de ce modèle mental, par exemple en favorisant les facteurs utiles pour le processus de construction, et en diminuant les facteurs qui le perturbent. Deux facteurs principaux ont été identifiés : la cohérence et le surcoût cognitif (Thüring, Hannemann *et al.*, 1995). Pour augmenter la cohérence d'un document, il faut fournir à l'utilisateur des repères l'aidant à identifier les composants majeurs du document et sa structure globale. Pour réduire l'effort mental de compréhension, il n'est pas suffisant d'avoir une structure globale cohérente. Il faut aussi communiquer cette structure à l'utilisateur. Il est alors possible de lui fournir les différents composants du document ainsi que leurs relations, par exemple à l'aide d'outils graphiques. Conklin caractérise le surcoût cognitif comme l'effort supplémentaire et la concentration nécessaire pour gérer plusieurs tâches en même temps (Conklin, 1987). Ce phénomène est dû aux capacités cognitives limitées du processus humain de traitement de l'information (Kaheneman, 1973). Dans un document hypermédia, ces efforts sont dus à l'orientation et la navigation. Pour éviter la désorientation, l'utilisateur doit non seulement connaître la structure globale du document, mais aussi ses déplacements dans cette

structure. Il doit en fait savoir où il est, d'où il vient et où il peut aller. Compréhension et orientation sont en effet très fortement liées. La compréhension d'un document hypermédia ou autre est donc liée à la compréhension de la structure globale de ce dernier, à sa présentation et à son contenu. Assurer une bonne compréhension d'une ressource complexe nécessite donc d'être capable d'adapter la structure globale d'une ressource, son contenu et sa présentation aux besoins utilisateurs.

Dans l'avenir, l'adaptation et la personnalisation seront des points clés pour l'utilisation, la vente et l'accès aux services et documents par les entreprises et les particuliers. Toutefois, il est important de faire la différence entre adaptable et adaptatif (Bra, 1999), même si la conjonction des deux est possible. Dans un système adaptable, l'utilisateur définit un ensemble de contraintes au travers d'une requête, d'un questionnaire ou d'un formulaire. Le système propose alors une ressource qui correspond au profil de l'utilisateur – les contraintes. Ce profil peut regrouper les préférences de l'utilisateur en terme de présentation (couleurs, type de média, style d'apprentissage, etc.) ainsi que son savoir-faire (qualifications, connaissances). Dans un système adaptatif, ce dernier observe le comportement de l'utilisateur et utilise ses observations afin d'adapter la présentation de l'information. L'évolution des préférences et de la connaissance de l'utilisateur est alors déduite (en partie) des accès aux pages Web. Le plus souvent, l'adaptation est obtenue au travers de la navigation de l'utilisateur, mais parfois le système peut aussi utiliser le comportement de l'ensemble des autres utilisateurs pour adapter la ressource. On se placera ici dans le cadre de l'adaptation/personnalisation (adaptable ou adaptatif) pour des utilisateurs au moyen de ressources liées à ces derniers – on ne se préoccupe pas de problèmes d'adaptation entre agents ou services qui ne font pas intervenir l'utilisateur ou des informations sur celui-ci.
Le cadre général que nous proposons est fondé sur le principe général suivant :

- Un ensemble de ressources est disponible sur le web (pages web, documents structurés, résultats de services, services) ;
- Un utilisateur « demande un résultat » ou pose une requête liée à sa tâche à un module logiciel en fixant un ensemble de contraintes qui sont des données non persistantes définies à partir d'un formulaire ou d'un simple hyperlien qui peut avoir été calculé pour ses besoins.
- Le module logiciel, à partir de cette demande, plus éventuellement un « profil » (données rémanentes caractérisant un utilisateur ou un groupe d'utilisateurs particuliers) sélectionne un sous-ensemble de ces ressources (requêtes logicielles). Un profil peut contenir des caractéristiques sur les connaissances, les préférences, les objectifs, les centres d'intérêts, etc. d'un utilisateur ou d'un groupe d'utilisateurs.
- Puis, le module logiciel **compose** ces ressources pour fournir à cet utilisateur un document ou service Web (page simple : liste ordonnée d'URL - adresses de documents, de page web ou de services - ou encore un site web ou service : portails, cours en ligne, musées virtuels, etc.), avec lequel l'utilisateur peut interagir à nouveau.

Dans un tel contexte, que peut faire le web sémantique dans le cadre de l'adaptation/personnalisation pour aider l'utilisateur ?

Une aide à l'utilisateur peut se faire aux étapes 2, 3 et 4 :

2) Demande utilisateur : proposer des termes adéquats au contexte de la tâche utilisateur, réécrire interactivement ou non sa demande en termes « reconnus » par les ressources et/ou les systèmes de gestion de ressources – hétérogènes, fournir à l'utilisateur un lien activant une requête adaptée à ses besoins.
3) Recherche et filtrage des ressources pertinentes à partir de la requête utilisateur et de son profil : la demande initiale et certaines données du profil utilisateur peuvent servir au filtrage des ressources pertinentes, que celui-ci soit

automatique ou interactif. L'enjeu est ici de réduire l'espace des ressources à celles qui s'apparient ou qui sont adéquates aux contraintes de la demande, aux besoins utilisateurs (via le profil).

4) Composition des ressources : après sélection des ressources pertinentes, il s'agit de les **organiser** pour les présenter à l'utilisateur. Cette organisation est le résultat de l'application de règles de composition qui peuvent prendre notamment les formes suivantes :
   - Une structure implicite résultat d'un calcul comme un simple tri – de type résultat de moteur de recherche, ordre temporel, etc.
   - Une structure explicite comme un modèle de tâche, une structure narrative, etc. qui organise un site Web (portail, musée virtuel, ...), etc.

   L'organisation sert de support pour la navigation et la compréhension. En effet, celle-ci est fortement liée à l'organisation et l'orientation dans un service ou document. Il est ici possible d'adapter/personnaliser l'organisation, le contenu, les modes d'interaction et la présentation des services ou documents.

1.2 EXEMPLES D'UTILISATION DE L'ADAPTATION / PERSONNALISATION DANS LE WEB SÉMANTIQUE

Un scénario d'utilisation du web sémantique et de l'adaptation articulé autour de l'enseignement à distance est présenté dans la partie dédiée aux annotations et aux méta-données. Ce scénario montre qu'il est possible d'effectuer de la recherche d'information adaptative à partir des méta-données qui caractérisent les ressources et d'un profil utilisateur. Par exemple à partir des connaissances acquises par un utilisateur dans un domaine particulier et des connaissances nécessaires à la compréhension d'une ressource, on peut juger de la pertinence de cette ressource et en informer l'utilisateur. Il est aussi possible de modifier l'organisation d'un cours, son contenu et sa présentation en fonction d'un profil utilisateur.

2 MÉTHODES, TECHNIQUES, OUTILS EXISTANTS SUR LESQUELS ON PEUT S'APPUYER

On a pu voir au travers du scénario que l'on doit être capable de calculer un document ou un service dynamiquement en fonction d'une spécification et/ou d'un profil utilisateur. L'adaptation/personnalisation dans le « web sémantique » peut s'appuyer sur les domaines suivants : les hypermédia adaptatifs, la modélisation utilisateur et les documents virtuels personnalisables (pour de plus de détails voir (Iksal, 2002)).

Le Web peut être vu comme un immense hypermédia dans lequel les informations sont reliées entre elles par des liens. Dès la fin des années 80, des équipes se sont intéressées à la notion d' « hypermédia flexible ». Un système hypermédia flexible est un système qui propose différentes vues d'un même hypermédia en fonction de l'utilisateur ou qui génère dynamiquement l'hypermédia et son contenu au moment de la consultation. Ils se déclinent principalement autour de deux approches (Oberlander & Milasavljevic, 1997) :

- L'utilisation des techniques venant de la recherche d'information et du filtrage afin d'évaluer la pertinence des nœuds d'un hypermédia et de modifier la présentation de l'information en fonction de cette évaluation.
- Les hypermédias adaptatifs (Brusilovsky, 1996) utilisent des caractéristiques de l'utilisateur afin de modifier la topologie d'hypermédias existants et de

présenter le contenu informationnel en fonction de ces caractéristiques[37]. Le point central de cette approche est la prise en compte de modèles utilisateurs.

2.1 LES HYPERMÉDIAS ADAPTATIFS

Le principe des hypermédia adaptatifs repose sur la constatation suivante : les hypermédia traditionnels présentent les mêmes pages et les mêmes hyperliens à tous les utilisateurs. Or, ceux-ci peuvent différer les uns des autres par leurs objectifs, leurs acquis, leurs antécédents et leurs connaissances par rapport au sujet traité par l'hypermédia. Par conséquent, ils ne seront pas intéressés par les mêmes informations de la page et ne suivront pas les mêmes liens lors de leur navigation. Les informations et liens non pertinents pour un utilisateur mais présentés à l'écran n'entraînent qu'une surcharge cognitive de celui-ci. L'utilisateur est facilement « perdu » dans l'étendue parfois vaste de l'hyperespace. Il a notamment souvent des difficultés à retrouver lui-même des chemins qu'il a précédemment suivis et il tourne facilement en rond dans son parcours des liens. Il éprouve souvent des difficultés à évaluer la quantité de pages qu'il a déjà visitées - ou qu'il lui reste à visiter. Rapidement, l'utilisateur peut mal interpréter le document, et se faire une mauvaise représentation mentale de la structure de l'hypermédia. Les objectifs des hypermédias adaptatifs sont donc de réduire la désorientation de l'utilisateur et le risque d'incompréhension du document qui en résulte le plus souvent.

Peter Brusilovsky nous en donne la définition suivante « By adaptive hypermedia systems we mean all hypertext and hypermedia systems which reflect some features of the user in the user model and apply this model to adapt various visible aspects of the system to the user » (Brusilovsky, 1996). Toutefois, il est important de connaître la différence entre un hypermédia adaptable et un hypermédia adaptatif (Bra, 1999) (cf. plus haut). Un hypermédia est composé de deux types d'éléments, des contenus et des liens. Les méthodes d'adaptation ont donc été classées en deux catégories, les méthodes de présentation adaptative qui manipulent le contenu des hypermédia et les méthodes de navigation adaptative qui concernent plus particulièrement les liens (Bra 1999 & Brusilovsky, 2001). L'idée de la présentation adaptative est d'adapter le contenu d'une page accédée par un utilisateur particulier, à sa connaissance, ses objectifs ou encore d'autres caractéristiques qui lui sont propres. La navigation adaptative a pour but d'aider les utilisateurs à trouver leur chemin dans l'espace d'information en adaptant la présentation des liens en fonction des objectifs, de la connaissance ou encore d'autres caractéristiques de l'utilisateur. Ces caractéristiques sont utilisées afin de classer les pages Web - appelées aussi nœuds - qui correspondent aux destinations des liens. Le système manipule alors les liens afin de guider l'utilisateur vers l'information la plus intéressante, la plus pertinente. Il se sert de la classe d'un nœud afin par exemple d'annoter ou de supprimer le lien qui pointe vers ce dernier.

[37] Il n'est pas ici question de générer dynamiquement l'hypermédia et son contenu mais uniquement d'effectuer de l'adaptation par rapport à un hypermédia préexistant.

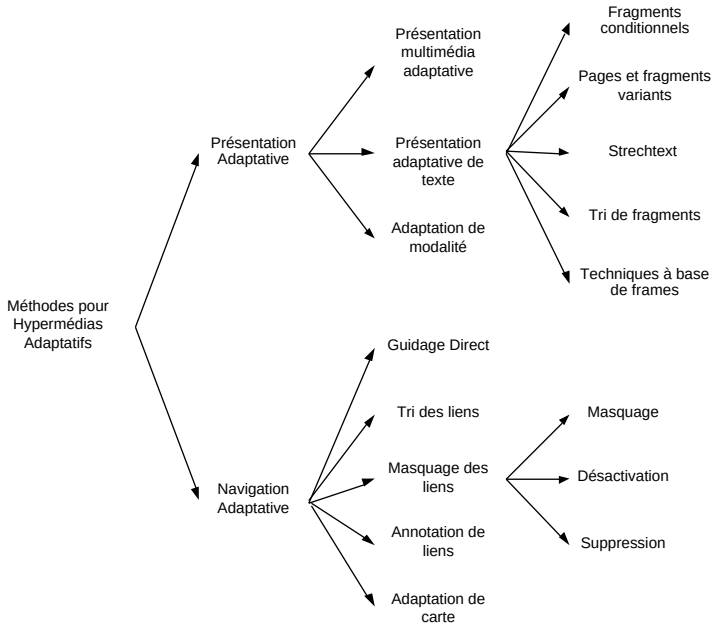

Fig1. Méthodes d'adaptation

Les travaux de recherche sur les hypermédia adaptatifs qui nous intéressent se sont principalement portés sur deux axes : l'adaptation de documents hypermédia (des sites Web – ELMART, AHA, INTERBOOK, etc. ou des applications locales) dont la structure globale était le plus souvent fixée à priori et sur la recherche d'information adaptative. Dans tous les cas, l'adaptation de la navigation, du contenu et des informations retrouvées nécessite des méthodes d'évaluation des liens, des contenus ou des informations retrouvées qui sont à la base des méthodes d'adaptation qui ont été proposées.

2.2 LA MODÉLISATION UTILISATEUR

Adapter, personnaliser un document ou une application pour un utilisateur particulier nécessite de disposer d'informations sur ce dernier permettant d'évaluer la pertinence des objets disponibles (fragments de texte, menus, boutons, etc.) ou d'aider le système à faire des choix. Le modèle utilisateur est « une source de connaissances, une base de données sur un utilisateur » (McTear, 1993). Plus précisément, il s'agit d'un ensemble de données persistantes 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.

Il existe différents types de modèles utilisateur qui peuvent être combinés. Le modèle individuel regroupe les caractéristiques propres à un individu. Il s'agit d'informations pouvant être soit renseignées par l'utilisateur, soit déduites par le système, soit acquises lors de l'utilisation de l'application. Le stéréotype correspond à un condensé des caractéristiques les plus représentatives d'un groupe - ou classe - d'individus, que nous pouvons qualifier de valeurs par défaut. Le stéréotype peut être utilisé tel quel, ou dans une optique d'individualisation de modèle (Rich, 1989). Le modèle de recouvrement associe pour un utilisateur particulier, une valeur à chaque concept du modèle de domaine. Chaque valeur correspond à une estimation du niveau de connaissance de l'utilisateur pour le concept. Le modèle de recouvrement est facile à mettre à jour mais souvent difficile à initialiser. Les actions de l'utilisateur sont analysées pour augmenter ou réduire le niveau de connaissance des concepts du

domaine (Brusilovsky, Eklund *et al.*, 1998). Le modèle partagé est un modèle réutilisable dans plusieurs applications. En effet, de nombreuses caractéristiques d'un utilisateur sont utilisées systématiquement par toutes les applications adaptatives. L'idée du modèle partagé est de disposer d'une partie commune et de parties spécifiques à chacune des applications ou tâches à réaliser. Kobsa et Wahlster (1989) proposent de disposer d'un modèle central et de plusieurs sous-modèles disjoints.
L'utilisateur peut être modélisé selon plusieurs points de vue :

- La connaissance de l'utilisateur apparaît comme la caractéristique la plus utilisée dans un système adaptatif. La façon la plus simple de gérer les connaissances est de mémoriser ce que l'utilisateur connaît ou ne connaît pas. Pour cela, on utilise souvent soit un stéréotype (modèle de groupe) soit un modèle de recouvrement (modèle individuel) qui associent aux concepts d'un modèle de domaine un niveau de connaissance (vrai/faux ou une valeur numérique).
- L'expérience et les compétences sont deux caractéristiques similaires à la connaissance définie dans le paragraphe précédent mais qui en diffèrent de par la nature même de l'information qu'elles représentent. L'expérience de l'utilisateur représente son savoir-faire, la familiarité et l'aisance qu'il possède avec le type de système qui lui est présenté. Les compétences possédées par l'utilisateur correspondent aux connaissances qui ne relèvent ni du domaine, ni de l'expérience mais qui sont néanmoins considérées comme pertinentes dans le fonctionnement du système.
- Les préférences de l'utilisateur peuvent porter sur une généralité et/ou un point particulier. Par exemple, un voyageur exige le côté fenêtre systématiquement, sauf pour l'avion, où il demande le couloir. Contrairement aux autres caractéristiques, les préférences ne peuvent être déduites par le système. En effet, chaque utilisateur possède ses propres préférences, et c'est à lui de les renseigner.
- Les objectifs : dans les systèmes d'information en ligne, les utilisateurs ont l'habitude de n'accéder qu'à une partie de l'espace d'information - ensemble de fragments - qui dépend principalement de leurs objectifs (Brusilovsky, Schwarz *et al.*, 1996; Vassileva, 1996). L'objectif est un état que l'utilisateur espère atteindre, et les plans décrivent les étapes pour y arriver. Les modèles de tâches sont très appropriés pour modéliser ces plans et par conséquent les objectifs des utilisateurs. Souvent ces démarches sont typiques d'une catégorie d'utilisateurs, c'est pourquoi on associe un modèle de tâches statique à un stéréotype (Garlatti, 1999). Ces modèles de tâches sont souvent utilisés en interaction homme/machine pour l'évaluation et l'interprétation (Tricot & Nanard, 1997) ainsi que la conception d'hypermédia (Fraissé 1997; Garlatti, Iksal *et al.*, 1999). Il n'est envisageable d'utiliser une telle méthode que lorsque les objectifs utilisateur ont une certaine stabilité.

Nous avons vu que l'adaptation nécessite des connaissances sur les utilisateurs. Ces connaissances sont obtenues grâce à la gestion de modèles utilisateurs qu'ils soient construits pour des communautés (stéréotypes) ou pour des individus. De nombreux travaux ont été menés sur les méthodes et outils de gestion de modèles utilisateurs. L'acquisition des caractéristiques utilisateur peut être réalisée par observation directe, interviews, questionnaires ou encore par apprentissage. Néanmoins dans le cadre du Web, le passage à l'échelle nécessitera d'avoir des méthodes adaptées.

2.3 LES DOCUMENTS VIRTUELS PERSONNALISABLES

Les hypermédias flexibles ont évolué vers la notion de document virtuel. Les documents virtuels sont nés d'un besoin d'interactivité et d'individualisation des documents, plus particulièrement sur le Web. 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, Vemuri *et al.*, 1996). En 1999, lors de la première conférence sur les documents virtuels, Carolyn Watters et Michael Shepherd (Watters & Shepherd, 1999) définissent brièvement le document virtuel comme étant un document qui ne possède pas d'état persistant et pour lequel toutes les instances sont générées lors de la consultation. Maria Milosavljevic (Milosavljevic, Vitali *et al.*, 1999) lors de l'introduction de la conférence, précise que le document virtuel est composé de l'information et des mécanismes nécessaires à la génération du document réel qui sera consulté par le lecteur. Un document virtuel possède donc les caractéristiques suivantes :

- réutilisation de différentes sources d'information,
- génération dynamique du document à la demande en fonction de l'utilisateur.

Il correspond à un espace d'information et des mécanismes de génération.

Si un utilisateur peut adapter/personnaliser le document réel produit en fonction de ses besoins, on parlera de *document virtuel personnalisable* (DVP), sinon uniquement de document virtuel. Un document virtuel peut être personnalisé à partir d'au moins un des trois critères suivants :

- Les contraintes utilisateur sont des données non persistantes définies à partir d'une requête, d'un formulaire ou d'un simple hyperlien et qui lui sont propres.
- Le modèle utilisateur : il s'agit d'un ensemble de données persistantes 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 (Rich, 1989; Brusilovsky, Schwarz *et al.*, 1996).

Le contexte (Brézillon, 2002) : support de lecture, taille d'écran, vitesse de transfert, conditions géographiques, météorologiques (Garlatti & Iksal, 2000), etc.

En 1999, Sylvie Ranwez et Michel Crampes 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, c'est-à-dire en respectant un modèle de l'utilisateur et des principes narratifs (Ranwez & Crampes, 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. L'indexation des fragments peut aller de l'utilisation d'un simple concept issu d'un modèle de domaine à celle plus complexe d'un schéma de méta-données. La caractérisation de ces fragments et de leur indexation ne peut pas être séparée de celle 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 d'un DVP, c'est l'ensemble des informations nécessaires au système pour composer le document. Cette étape dépend bien évidemment des principes mis en oeuvre 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 dernières ne sont pas toujours accessibles aux auteurs. 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 ; elles ne peuvent être conçues l'une sans l'autre et forment un tout indissociable.

Dans le cadre des documents virtuels personnalisables sur Internet qui génèrent des hypermédias, il est intéressant de prendre en compte les méthodes et techniques

d'adaptation proposées dans le domaine des hypermédias adaptatifs. La personnalisation, l'adaptation pour un utilisateur donné est favorisée par l'utilisation des modèles utilisateurs qui représentent ces utilisateurs. En ce qui nous concerne, les DVP sont fondés sur des informations provenant de sources hétérogènes, la réutilisation et le partage sont des enjeux des DVP qui peuvent être assurés par le Web Sémantique.

3 TRAVAUX ET RÉSULTATS EXISTANTS DU WEB SÉMANTIQUE

Les travaux actuels sur l'adaptation dans le Web sémantique peuvent être décomposés en deux parties : l'une sur les hypermédias adaptatifs et l'autre sur les documents virtuels adaptatifs. Dans le domaine des hypermédias adaptatifs de nombreux travaux sont en cours sur l'adaptation au support physique et aux préférences de l'utilisateur. Le système Cuypers a pour objet la génération automatique de présentations hypermédia temporelles adaptées aux préférences et capacités des utilisateurs et du support physique (Ossenbruggen, Hardman *et al.*, 2001). La méthodologie Hera a été utilisée pour la conception de systèmes d'information intelligents et adaptatifs sur Internet. L'adaptation se fait par rapport au support physique et à l'historique de navigation des utilisateurs (Frassincar & Houben, 2002). Des travaux sont également en cours sur des systèmes adaptatifs à l'initiative de l'utilisateur (Aragão, Fernandes *et al.,* 2001). Dans le domaine des documents virtuels personnalisables, on peut citer des travaux sur la création de mémoires d'entreprise (Fortier & Kassel, 2002), ainsi que le projet RNRT KMP (Knowledge Management Portal), les dossiers thématiques personnalisables (Iksal & Garlatti 2001; Iksal, Garlatti *et al.*, 2001) (projet ICCARS) (Iksal, 2002) et le E-learning (Projet Candle, http://www.candle.eu.org/). Dans la cadre des prochains workshops intitulé Adaptive Hypermedia and Adaptive Web-Based Systems (http://wwwis.win.tue.nl/ah2003/) d'autres travaux seront présentés.
Il existe actuellement de nombreux travaux sur la description des politiques de gestion des données utilisateur – publiques/privées : travaux du W3C autour de P3P (http://www.w3c.org/P3P/) (possibilité de comprendre et de gérer l'utilisation de ses données personnelles).

4 RECHERCHES FUTURES POUR LE WEB SÉMANTIQUE

Les futures recherches sur le Web Sémantique pour l'adaptation/personnalisation peuvent s'articuler autour des cinq points suivants :

- la modélisation utilisateur et des utilisations,
- la protection de la vie privée,
- la recherche d'information,
- les services et documents adaptatifs,
- l'ergonomie des services et des documents.

4.1 MODÉLISATION DES UTILISATEURS ET DES UTILISATIONS

Pour la modélisation utilisateur/utilisation, il est important d'avoir des standards de modèles utilisateur/utilisation disponibles sous formes d'ontologies pour des catégories de services/documents (possibilité d'enrichissement dynamique du modèle utilisateur courant afin qu'il puisse bénéficier de services adaptatifs/personnalisables).

En effet, si nous pouvons espérer dans l'avenir disposer de nombreux services et/ou documents adaptatifs, il sera nécessaire d'avoir accès à différentes catégories de modèles utilisateur/utilisation qui ne seront sans doute pas sans intersection commune. Il en de même aujourd'hui pour les méta-données puisque différents schémas de méta-données ont des champs qui sont communs et d'autres qui diffèrent. Il sera donc nécessaire de standardiser certains modèles.

Actuellement, la plupart des méthodes d'analyse du comportement des utilisateurs sur le Web pose le problème du passage à l'échelle et du choix des critères pertinents à analyser. En fait, on peut disposer d'énormes quantités de données, mais il est difficile de savoir comment les analyser et lesquelles sont pertinentes pour un objectif donné des utilisateurs et/ou des concepteurs. Il y a donc un réel problème pour l'acquisition et la maintenance des modèles utilisateur – dynamiques et/ou statiques, ainsi que pour l'acquisition de connaissances sur les utilisateurs/utilisations par observation de comportements. Il s'agit notamment ici de développer des observatoires des usages, et de mettre en place des outils permettant de saisir des *utilisations* significatives, et non de simples « logs ininterprétables ». A partir de ces observatoires, il faudrait pouvoir déterminer les paramètres pertinents à observer chez les utilisateurs par classe de services et en déduire des standards de modèles utilisateur qui seraient réutilisables et partageables.

4.2 PROTECTION DE LA VIE PRIVÉE

Comme on vient de le voir, l'utilisation de modèles utilisateur est des points clés de l'adaptation/personnalisation. Chacun aura un ou plusieurs modèles qui seront mis à jour par les services accédés. Mais qui va posséder/archiver ses modèles ? Comment peut-on contrôler l'utilisation de ceux-ci *a fortiori* lorsque certains logiciels feront des déductions à partir des données enregistrées ? Il faut donc mettre en place des politiques claires sur l'utilisation, l'archivage et la mise à jour des modèles. Il faudra établir des standards de services et de qualité de services à propos de la sécurité et du partage entre données privées/publiques. Il est probable qu'il faudra développer des serveurs de modèles utilisateurs indépendants des services proposés afin d'assurer un minimum d'indépendance entre l'archivage, l'utilisation, la mise à jour et les fournisseurs de services.

Il serait donc intéressant d'avoir des services muni de « certificats » qui assurent une qualité de service et de confidentialité conforme à des standards établis. Chaque classe de service aura des méta-données conformes à leur « certification » et qui seront ainsi accessible par des moteurs de recherche utilisant les schémas de méta-données adéquats qui devront être à disposition sur internet.

4.3 LA RECHERCHE D'INFORMATION

Dans l'avenir, il existera très certainement plusieurs ontologies sur le même domaine ou ayant le même objectif. Elles vont devoir coexister et collaborer. En effet, il est difficile de mettre tout d'accord avec une unique ontologie dans certains domaines. Il sera donc nécessaire d'aider les utilisateurs à poser une requête par sélection d'une ou plusieurs ontologies pour un domaine donné ou par proposition de différents vocabulaires. Il en est de même pour la transformation d'une requête en une autre fonction des ressources hétérogènes disponibles sur le réseau et de leur modélisation. Il s'agit ici de développer des systèmes médiateurs capables d'établir les correspondances entre différentes modélisation d'un domaine.

Pour la recherche d'information adaptative, Il serait également important d'être capable de filtrer des ressources et/ou d'annoter et/ou de classer des ressources en fonction de modèles utilisateur (situation géographique, tâche courante, préférences, support physique ou nature du réseau de transmission). Pour cela, il faut développer

des méthodes d'évaluation des ressources qui devront prendre en compte des schémas de méta-donnés, des modèles utilisateur et des méthodes d'adaptation qui devront tous être « standardisés » - ontologies pour les schémas de méta-données, les modèles utilisateurs et les modèles d'adaptation.

4.4 SERVICES ET DOCUMENTS ADAPTATIFS /PERSONNALISABLES

Pour fournir des services et documents adaptatifs aux utilisateurs, il est nécessaire d'être capable de construire dynamiquement ces services et ces documents à partir de ressources, d'une spécification du service ou document, d'un modèle utilisateur et de méthodes d'adaptation et d'évaluation des ressources. Du point de vue des fournisseurs de services ou documents, il est important d'avoir des méthodes de création et de maintenance, les plus automatisables possible. Pour cela, il faut donc étudier l'apport des ontologies, des méta-données et des mécanismes d'inférence qui assurent cette automatisation.

Pour l'adaptation, les relations entre les méta-données, la granularité et les modèles utilisateurs jouent un rôle important. Par exemple, l'adaptation se fait-elle par le choix – filtrage - d'un document ou service pertinent ou par la sélection d'une partie du service ou document, certains types de méta-données sont-elles dédiés à l'adaptation ? Est-ce un invariant pour tous les services ou documents ou dédiés à certaines classes de services ou documents ? etc. La standardisation et/ou la description sémantique de ces relations entre modèles utilisateurs, méta-données, granularité et adaptation devrait assurer une bonne réutilisation de ces services et documents adaptatifs.

Du coté des auteurs/concepteurs, il est nécessaire de développer des environnements qui permettent autant que possible de voir le système auteur comme une « Glass Box » afin que ceux-ci soient en mesure de connaître le comportement du système à partir de leur spécification et plus particulièrement pour l'adaptation. En effet, un système peut potentiellement engendrer une infinité de services ou documents réels. Pour cela, il faut analyser en quoi l'utilisation des ontologies et du raisonnement peu faciliter les tâches des auteurs et quelles sont les IHM les plus pertinents et comment pourrait-on les engendrer dynamiquement à partir de leur description à un niveau sémantique. IL est en effet tout à fait possible d'utiliser les principes des DVP pour la création d'interface adaptative. Là encore, l'utilisation de description au niveau sémantique devrait permettre une mise en œuvre plus rapide et une meilleure réutilisation.

Pour chaque service, ses politiques d'adaptation et les modèles utilisateur requis doivent être clairement identifié afin qu'un modèle « global » de l'utilisateur soit mis à jour sans dupliquer certaines données déjà présentes et non spécifiques à un service. Il faut donc pour cela établir des schémas de méta-données particuliers et/ou des catégories de schémas.

4.5 ERGONOMIE DES SERVICES ET DOCUMENTS

Pour l'ergonomie des services et des documents, il est nécessaire de développer des recherches dans les domaines de la sociologie, de la psychologie cognitive et de l'ergonomie cognitive sur les aspects suivants :

- Évaluation de l'utilisabilité, de l'utilité, de l'acceptabilité et de l'ergonomie des services et documents.
- Études sur la présentation de l'information, les formats et la compréhension, …

RÉFÉRENCES

ARAGÃO V. R., FERNANDES A. A. A. et al. (2001). Towards an Architecture for Personalization and Adaptivity in the Semantic Web. *Proc. 3rd. Intl. Conf. on Information Integration & Web-Based Applications & Service, IIWAS 2001,* Linz, Austria. Austrian Computer Society: 139-149.

BRA P. D. (1999). Design issues in adaptive hypermedia application development. *Workshop on Adaptive Systems and User modeling on the World Wide Web,* Toronto & Banff.

BRÉZILLON P. (2002). Hors du contexte, point de salut. S. O. Communicants". Autrans, France.

BRUSILOVSKY P. (1996). "Methods and techniques of adaptive hypermedia." *User Modeling and User-Adapted Interaction* **6**(2-3): 87-129.

BRUSILOVSKY P. (2001). "Adaptive Hypermedia." *User Modeling and User-Adapted Interaction* **11**: 87-110.

BRUSILOVSKY P., EKLUND L. et al. (1998). Web-based Education for all: A tool for development adaptive courseware. *Seventh International World Wide Web Conference. Computer Networks and ISDN Systems,* (30): 291-300.

BRUSILOVSKY P., SCHWARZ E. et al. (1996). ELM-ART: An intelligent tutoring system on World Wide Web. *Third International Conference on Intelligent Tutoring Systems, ITS-96,* Montreal: 261-269.

CONKLIN J. (1987). "Hypertext: An introduction & Survey." *IEEE Computer,* 20(9):17--41

FORTIER J. Y. & KASSEL G. (2002). Génération de documents virtuels personnalisés à partir de modèles de connaissances. *Documents Virtuels Personnables 2002,* Brest, 10-11 Juillet, S. Garlatti, M. Crampes, ENST Bretagne: 115-126.

FRAISSÉ S. (1997). Une approche systèmatique pour la conception de documents hypermédias. *Hypertextes et Hypermédia. J. P. Balpe, A. Lelu, M. Nanard & I. Saleh, Hermes:* 249-260.

FRASSINCAR F. & HOUBEN G.-J. (2002). Hypermédia Presentation Adaptation on the Semantic Web. *Adaptive Hypermedia and Adaptive Web-Based Systems,* Malaga, Springer Verlag, *Lecture Notes in Computer Science*: 133-142.

GARLATTI S. (1999). Adaptive Web Server for On-line Information Systems. *5th European Summer School, EUNICE'99,* Barcelone.

GARLATTI S. & IKSAL S. (2000). Concept Filtering & Spatial Filtering in an Adaptive Information System. *Adaptive Hypermedia & Adaptive Web-Based Systems. P. Brusilovsky, O.* Stock & C. Strapparava. Trento, Springer Verlag**:** 315-318.

GARLATTI S., IKSAL S. et al. (1999). Adaptive On-line Information System by means of a Task Model & Spatial Views. *Second Workshop on Adaptive Systems & User Modeling on the World Wide Web,* Toronto & Banff, Eindhoven University of Technology.

GRUBER T. R., VEMURI S. et al. (1996). Model-based Virtual Document Generation, *Knowledge Systems Laboratory,* KSL-96-16.

IKSAL S. (2002). Spécification Déclarative et composition sémantique pour des documents virtuels personnalisables. Thèse IASC. Brest, ENST Bretagne.

IKSAL S. & GARLATTI S. (2001). Revisiting and Versioning in Virtual Special Reports. Hypermedia: Openness, Structural Awareness and Adaptivity, *Third Workshop on Adaptive Hypertext & Hypermedia. S. Reich, M. M. Tzagarakis and P. D. Bra.* Åarhus, Denmark, Springer Verlag**:** 264-279.

IKSAL S., GARLATTI S. et al. (2001). Semantic composition of special reports on the Web: A cognitive approach. *Hypertextes and Hypermédia H2PTM'01. P. Balpe, S. Lelu-Merviel, I. Saleh & J.-M. Laubin,* Hermès**:** 363-378.

KAHENEMAN D. (1973). *Attention & Effort.* Englewood Cliffs, Prentice Hall.

KOBSA A. & WAHLSTER W. (1989). *User Models in Dialog Systems,* Springer Verlag.

MCTEAR M. F. (1993). "User modelling for adaptive computer systems: a survey of recent developments." *Artificial Intelligence Review* **7**: 157-184.

MILOSAVLJEVIC M., VITALI F. et al. (1999). Introduction of the virtual document workshop. *Workshop on Virtual Document, Hypertext Functionality & the Web*, Toronto.

OBERLANDER J. & MILASAVLJEVIC M. (1997). Flexible Hypertext - introduction. *Hypertext'97, Proceedings of Flexible Hypertext Workshop,* Southampton, UK.

OSSENBRUGGEN J. R. V., HARDMAN H. L. et al. (2001). Hypermedia and the semantic web: a research agenda, Centrum voor Wiskunde in Informatica.

RANWEZ S. & CRAMPES M. (1999). Conceptual document and hypertext documents are two different forms of virtual document. *Workshop on Virtual Document, Hypertext Functionality and the Web,* Toronto.

RICH E. (1989). Stereotypes and user modeling. user models in dialog systems. A. Kobsa & W. Wahlster. Berlin, Springer verlag**:** 35-51.

THÜRING M., HANNEMANN J. et al. (1995). "Hypermedia & Cognition: Designing for Comprehension." *Communication of the ACM* **38**(8): 57-66.

TRICOT A. & NANARD J. (1997). Un point sur la modélisation des tâches de recherche d'informations dans le domaine des hypermédias. *Hypertextes and Hypermédia. J. P. Balpe, A. Lelu, M. Nanrd & I. Saleh,* Hermes**:** 35-56.

VASSILEVA J. (1996). "A Task-Centered Approach for user Modeling in a hypermedia Office Documentation system." *User Models & User Adapted Interaction* (6): 185-223.

WATTERS C. & SHEPHERD M. (1999). Research issues for virtual documents. *Workshop on Virtual Document, Hypertext Functionality & the Web*, Toronto.