

Sophie Chauvin
Univ. Lille 3, GERICO

Fabrice Papy
Univ. Paris 8, DnU

Mohamed Sidir
Univ. de Picardie, SASO

Peter Stockinger
INALCO, FMSH, EsCOM

Le portail institutionnel Persée à l'épreuve des usages : croiser les approches méthodologiques en Sciences Humaines pour améliorer le partage de connaissances scientifiques en libre accès

Résumé : Les recherches antérieures spécifiques menées par 4 enseignants-chercheurs en Sciences de l'Information et en Sémiotique trouvent un terrain de convergences dans une étude des usages du portail institutionnel Persée. Ergonomie visuelle, organisation documentaire hypertextuelle, sémiotique du dispositif et communication instrumentée constituent autant des aspects saillants par lesquels l'étude est menée et qui accompagnent la mise en œuvre de la nouvelle version du portail.

Abstract: The specific previous studies conducted by 4 teacher-searchers in Information Science and Semiotics find a convergence ground with a study on uses of the institutional portal Persée. Visual ergonomics, hypertext organization, disposition of the semiotics and instrumental communication are the key aspects of this study and form the base of the implementation of the portal new version.

Introduction

Le portail documentaire Persée (<http://www.persee.fr>) créé par le Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche français, se présente comme un programme de diffusion élargie de corpus scientifiques provenant de collections rétrospectives de revues scientifiques de référence en sciences sociales et humaines. Depuis le 21 janvier 2005, ce sont plusieurs milliers d'articles, issus d'une vingtaine de revues, qui sont mis gratuitement à la libre consultation d'un large public, chercheurs ou non, à partir d'un portail dont l'objectif exprimé est d'apporter une exploitation enrichie de ces revues en termes de recherche et de consultation.

Ce dispositif qui s'inscrit dans un processus cumulatif de numérisation a abouti à la mise en place d'une architecture d'une extrême complexité compte tenu de ses dimensions éditoriales (contrats inédits auprès des revues et éditeurs afin de garantir la commercialisation de la production courante), organisationnelles (chaîne de production intégrée couvrant de la numérisation à la mise en ligne des documents) et technologiques (technologies XML, développements open source, interopérabilité avec des sites partenaires et les Environnements Numériques de Travail des Universités).

Après 2,5 années d'existence, malgré une augmentation très significative du nombre de revues depuis la mise en ligne (triplement) et plusieurs centaines de milliers de connexions, la réussite du portail en termes d'appropriation par ses usagers reste mitigée.

La difficulté à lire significativement une régularité dans la consultation des utilisateurs inscrits, l'inutilisation des forums (générique et thématique) intégrés au portail, le peu d'intérêt des internautes pour les services complémentaires (pourtant gratuits), le peu d'utilisation des fonctionnalités comme la recherche avancée et, finalement la difficulté à cerner les profils des visiteurs et identifier leur utilisation effective du portail sont autant de contours flous qui rendent indiscernable l'impact tangible du portail sur la communauté des chercheurs.

Ces constats de succès relatif tout en révélant l'« action dans l'incompréhension » des usagers à propos de la manipulation de dispositifs informatiques, œuvres de concepteurs et d'informaticiens chevronnés, soulignent dans le même temps, la rupture qui existe entre l'appréciation des concepteurs sur des dispositifs qui selon eux "fonctionnent correctement" et l'appréciation des utilisateurs de ces mêmes dispositifs (Folcher 2005) (Papy Folcher Sidir Pargam 2004) (Bruillard 2000) (Jeannert 2000) (Pavé 1989) (Corson 1982). Pour eux, la difficulté majeure porte sur le rapprochement de leur activité cognitive avec les fonctionnalités techniques dont sont assortis les dispositifs qu'ils utilisent.

Quand la logique « projet » l'emporte sur la démarche « usage »

C'est dans le cadre de la refonte programmée du portail Persée que notre équipe, composée de quatre enseignants-chercheurs en Sciences Humaines spécialisés sur les problématiques des usages des TIC, s'est vue confiée depuis septembre 2006 une étude analytique et prospective, sur l'adéquation des fonctionnalités du portail avec les attentes et les besoins de la communauté des chercheurs en SHS.

Il est important de noter que la mise en place du portail Persée n'a pas bénéficié d'une analyse systématique et rigoureuse des usages *a priori* à la hauteur de ce que les chercheurs sont en mesure de produire (Chauvin 2005) (Ciaccia 2005) (Jannes-Ober 2005) (Dinet 2002) (Le Marec 1990) .

Préalablement, lors de la phase de conception du portail PERSEE, un état de l'art des portails existant et proposant des contenus scientifiques en ligne a été dressé. L'objet de ce recensement portait sur l'étude des services et des fonctionnalités proposés et la manière dont ils étaient mis à disposition. Cette analyse a porté sur différents sites classés en :

- sites de diffusion de contenu patrimonial (documents en mode image) ;
- sites d'édition électronique de revues scientifiques en sciences humaines et sociales ;
- sites d'édition électronique de revues scientifiques en sciences, technique et médecine.

Cette première approche avait été complétée par une évaluation d'une sélection de sites diffusant l'information scientifique et des maquettes du site PERSEE par des groupes d'utilisateurs (étudiants, chercheurs, professionnels de la documentation). Sur certains

des éléments fonctionnels de la maquette, placés en situation de recherche, les utilisateurs devaient trouver des documents, lancer des requêtes, conserver leurs résultats de recherche, imprimer un document etc. Cette méthode présentait l'avantage d'observer des personnes en situation d'utilisation du portail PERSEE et in fine de valider par l'« usage », la pertinence des fonctionnalités proposées et quelques éléments d'ergonomie.

Les orientations de conception dont l'actuelle version a héritées proviennent pour l'essentiel des expériences de l'équipe réalisatrice PERSEE en matière de développement de sites Internet. L'accent a été mis sur le fait de :

- faciliter la navigation des utilisateurs en privilégiant une interface sobre ;
- permettre aux utilisateurs quel que soit leurs profils d'accéder de la même façon à l'information ;
- permettre *en moins de trois clics* d'accéder à l'information recherchée ;
- présenter dès la page d'accueil le contenu du site et ce que l'on va pouvoir y trouver.

Ces conditions de conception ont influencé un contexte d'élaboration du portail Persée où des logiques essentiellement technologiques et politiques ont cadencé le projet institutionnel techno-informatique.

L'étude d'usages actuellement menée vise d'une part à revisiter les fonctionnalités disponibles aujourd'hui sur Persée (et tenter d'identifier la raison de l'inutilisation de certaines d'entre elles) et d'autre part à anticiper les impacts sur les usages que les extensions du portail en matière de contenus et de fonctionnalités ne manqueront pas d'avoir (relations avec des sites partenaires).

L'utilisabilité du site en correspondance avec les objectifs déclarés (<http://www.persee.fr> => Présentation du portail) et corrélés à la capacité des utilisateurs/usagers à se l'approprier à la hauteur du fonds documentaire mis en ligne et des fonctionnalités techniques proposées ne relève plus uniquement de la compétence des concepteurs et développeurs impliqués.

En effet, la mise en œuvre de tels artefacts introduit l'obligation de s'adapter à des contraintes technologiques incompressibles qui éloignent – presque naturellement – l'équipe de conception-réalisation des attentes sociales, culturelles et sémio-cognitives des usagers potentiels, attentes qui sont peu explicitées et rarement formalisées (Pavé 1989).

En l'absence d'indications précises et exhaustives, c'est bien la figure de l'utilisateur du dispositif ("le profil technique et informatif") qui s'avère complexe à appréhender et que l'équipe de conception-réalisation va imaginer à partir d'une représentation intuitive et personnelle d'un usager-type auxquels ils affecteront hypothétiquement tel contexte socio-cognitif, tels objectifs de consultation, telles habiletés techniques, etc.

Par exemple, la possibilité d'une "recherche avancée" dans la plupart des systèmes de recherche d'informations (moteur de recherches, annuaires, catalogues, portail, etc.) qui offre aux utilisateurs de composer des équations de recherches complexes (ET, OU, SAUF) souligne ce décalage entre la représentation que les concepteurs peuvent avoir de l'habileté technique des utilisateurs et leur savoir-faire réel (Jannès-Ober, 2005)

(Simonnot 2002).

"...Des études récentes montrent que, dans la majorité des cas, les requêtes formulées comportent deux termes ou moins (Jansen, Spink, Saracevic, 2000), ce qui est bien peu pour représenter un besoin d'information. (...) On sait que la formulation d'une requête à l'aide d'opérateurs booléens est contre-intuitive pour les utilisateurs non spécifiquement formés à cette démarche (Avrahami et Kareev, 1993)..." (Dinet 2002)

Fonctionnalités actuelles du Portail Persée

Le portail fonctionne actuellement sur deux principes : navigation dans les corpus numérisés et recherche. Un ensemble d'outils et de services a été développé pour permettre une exploitation et une appropriation de l'information comparables aux sites de diffusion courante :

Services pour l'utilisateur non enregistré :

- Recherche, qui peut porter sur les mots du titre, le nom de l'auteur et le texte intégral ;
- Navigation au sein des numéros de revues, facilitée par des sommaires et des tables des matières numériques ;
- Accès aux documents en mode image ainsi qu'à l'OCR non corrigé ;
- Impression des articles ;
- Téléchargement des articles ;
- Accès à des références bibliographiques enrichies de liens vers les résumés, lorsqu'ils existent ;

Services supplémentaires pour l'utilisateur enregistré (espace personnel) :

- Sauvegarde des termes de recherche ;
- Alerte par un système de DSI quand un article correspondant aux termes de recherche enregistrés est ajouté à la base PERSEE ;
- Historique des recherches ;
- Panier des articles sauvegardés ;
- Configuration du nombre de résultats par page ;
- Export des références bibliographiques ;

Autres services :

- Forum du portail PERSEE ;
- Un forum spécifique pour chaque revue ;

Approches croisées et méthodologies complémentaires

De par leurs travaux de recherches respectifs, les membres de l'équipe de recherches en Sciences Humaines concernés par cette étude, entretiennent un rapport étroit avec les Technologies de l'Information et de la Communication en matière de conception et d'analyse de l'activité instrumentée où l'approche « conception dans l'usage » demeure étroitement liée à celle de la « conception pour l'usage » (Folcher et Rabardel 2004) : archives audiovisuelles numériques (AAR), plates-formes d'enseignements à distance (INES), dispositifs de communications synchrones et asynchrones (INES), agrégateurs de contenus (Web Feedback, K-Web Organizer), catalogues en ligne de bibliothèques (Visual...Catalog), moteur de recherches (NeuroWeb), etc.

Pour cette étude sur le portail Persée, trois modes d'analyses ont été retenus :

- 1) Monosémie et polysémie des fonctionnalités documentaires et techniques de Persée (Risques de désorientation dans la navigation et d'incompréhension de la cohérence globale du dispositif. Analyse a priori des critères ergonomiques).
- 2) Approche sémiotique par les genres (structures scénarielles sous-tendant l'organisation du site et leurs impacts sur les types de consultation).
- 3) Analyse des usages par l'analyse des traces de communications asynchrones via les forums.

Monosémie et polysémie des fonctionnalités documentaires et techniques de Persée

Les travaux expérimentaux sur la bibliothèque et les interfaces sont à l'origine de questionnements déjà anciens sur l'accessibilité introduite par les TIC et largement reprise par les systèmes de recherche d'informations (SRI) (Chauvin 2006) (Chauvin 2005).

En effet, l'accessibilité documentaire ne peut être réduite à la seule dimension technique de publication/diffusion des ressources nouvellement numérisées. Une ressource numérique (ou numérisée) ne devient pas immédiatement accessible par sa seule publication au sein d'un système de gestion de contenus.

En référence aux approches de Bernard Miège sur la communication personnes-systèmes d'information, l'accessibilité dont il est question ici est la résultante de l'amplification produite par les TIC (notamment dans les SRI) qui portent à l'appréciation des usagers une multitude de parcours et de relations de sens. Une accessibilité sémio-cognitive aux ressources/documents ne peut advenir sans l'accompagnement des usagers par un environnement de cohérence à la fois technique et structurelle : "...certaines expériences (...) montrent qu'avec l'expérience et la pratique, un individu modifie ses stratégies de recherche d'informations puisqu'il connaît de mieux en mieux la structure du corpus informationnel, le contenu informationnel et les outils dont il dispose...". (Dinet, 2002, 134)

Cependant cette cognition située nécessite de saisir la cohérence et le sens de l'organisation des données, d'appréhender les données proposées à la consultation et de se représenter mentalement leurs places dans la structure d'ensemble. Si cette représentation n'est pas claire, c'est tout le processus de recherche d'information (RI) qui est remis en cause (Baccino, 2004) (Ciaccia, 2005).

Des travaux déjà anciens sur les systèmes hypertextes ([Papy 95] [Brown 88a] [Nielsen 90] [Conklin 87] [Foss 89]) ont mis en évidence que l'accessibilité technique

tout en offrant des perspectives originales en matière de consultation, était, par son caractère d'immédiateté, source de désorientation. A ce sujet, trois causes possibles ont été identifiées :

- la navigation dans l'espace logiciel : le lecteur ne comprend pas les facilités de navigation que lui offre le système technique ;
- la navigation dans l'espace conceptuel : le lecteur ne comprend pas comment l'auteur a organisé l'information pour la représenter sous une forme hypertextuelle (par exemple comment il a défini la hiérarchie et comment il l'a enrichi par les liens supplémentaires). La perception qu'il peut avoir de la manière dont l'information est organisée peut alors entrer en collision avec le modèle de conception de l'auteur et ainsi expliquer des activations de liens inadéquates.
- la navigation dans le texte : le lecteur se perd de la même manière qu'un automobiliste se perd en prenant successivement des routes à droite et à gauche en espérant reprendre son itinéraire initial. Le lecteur, sans cesse sollicité par des liens à activer, se laisse tenter par des digressions, qui lui font perdre le fil de sa navigation de départ. Les multiples décrochements qu'il a effectués finissent par le laisser à un point du document qui ne correspond plus à sa recherche/lecture première.

Le troisième objectif des concepteurs de Persée est de « *permettre une exploitation enrichie de ces revues en termes de recherche et de consultation* ».

C'est cet objectif que nous rapprochons de la question de l'accessibilité et des relations que les modalités fonctionnelles disponibles sur Persée entretiennent avec elle.

Par « Permettre une exploitation », nous entendons permettre aux usagers d'utiliser pleinement cette mise en ligne de documents en s'appuyant sur l'ensemble des orthèses, extensions et fonctionnalités qu'offre le dispositif technique.

Neutraliser la question du "profil" de l'utilisateur

Nous considérons d'emblée qu'il n'est pas envisageable d'obtenir des profils identifiés des usagers du portail, bien que celui-ci soit destiné aux chercheurs en SHS, le libre accès au portail (dans le sens où il est gratuit) laisse entrevoir un public potentiel beaucoup plus large dont les caractéristiques, les motivations et les centres d'intérêts seraient d'une autre nature que celle relevant du monde de l'enseignement supérieur et de la recherche.

Les incertitudes majeures relatives à l'aisance technique des usagers avec les outils informatiques (et la consultation Web) rendent cette approche "profil" particulièrement hasardeuse.

Le portail Persée se trouve à la convergence d'un environnement technique autonome (la plate-forme technologique d'accueil est dotée de fonctionnalités opératoires qui demeurent valides indépendamment du contenu) et de ressources documentaires (revues papier) dont l'existence et l'exploitation étaient isolées de toute interaction informatique.

On peut considérer a priori que les usagers potentiels ont un intérêt manifeste pour la nature du fonds et que la médiation technologique qui leur est proposée vise à mieux

l'exploiter (recherches sur une revue à des recherches transversales impliquant plusieurs revues).

Ces deux aspects ; documentaire et technologique, vont constamment interférer dans l'utilisation du dispositif. Notre étude portera précisément à identifier les interactions possibles qui relèvent d'une part du paradigme technique et d'autre part du paradigme du contenu.

Le paradigme de la technique :

Les différentes modalités d'interactions que propose Persée (menus , icônes, "hotwords", etc.) vont orienter l'usager-utilisateur vers des manipulations para-textuelles (la recherche simple ou avancée, les forums, les préférences, le compte usager, les requêtes ayant déjà été tapées, les articles retenus, le panier, etc.) qui viennent s'ajouter aux manipulations du navigateur qu'il utilise habituellement et qui déterminent des fonctionnalités techniques spécifiques (précédent, suivant, historique, onglet, taille du texte, enregistrement de signets, etc.)

Le paradigme du contenu

Des modalités d'interactions spécifiquement liées au contenu vont solliciter l'usager vers une activité de lecture (accéder à une revue, accéder à une année, un numéro, accéder au sommaire, passer à l'article suivant, se rendre à une page donnée, représentation en vignette, exportation en PDF pour impression, page suivante, page précédente, etc.).

C'est à partir de la notion générique d'utilisabilité (*usability*) appliquée à l'évaluation des interfaces de systèmes de recherche d'information que nous aborderons l'étude du portail Persée. Cette approche à pour prolongement le principe de conception dans l'usage et pour l'usage (Henderson 1991) (Dittrich 2002) par l'observation et l'analyse des usages qui résultent de l'interaction au dispositif et dans le retour qui en est fait pour la conception.

C'est une boucle rétroactive où conception et évaluation s'enrichissent mutuellement par l'analyse qui en découle (cf. figure 1).

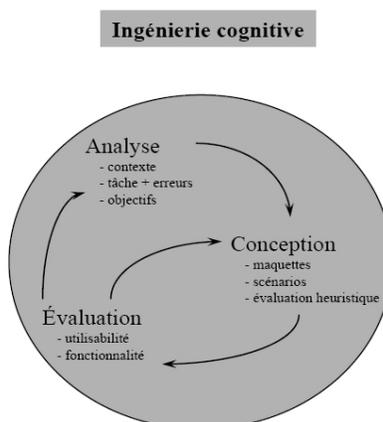


Figure 1 : Schéma du groupe IPSI du CRIM sur l'"intégration avec méthode de développement de logiciel". L'évaluation de dispositifs à partir d'observations et de prélèvements de données se prolonge vers l'analyse qui en découle. La synthèse de cette analyse est ramenée vers la conception et l'ensemble des étapes forme une boucle qui intègre le principe du *design-in-use* / *design-for-use*.

Méthodologie

L'observation en amont de Persée visera à distinguer toutes les interactions possibles et à les identifier en fonction de leur nature technique ou documentaire et de leur caractère sémique. Une mauvaise identification des interactions techniques et documentaires est selon nous, susceptible de produire une forme d'inaccessibilité constituée de désorientations, d'incompréhensions et de ruptures dans le parcours de lecture de l'utilisateur.

Afin d'identifier et isoler sur le site les interactions notamment ambiguës, qui relèvent du monosémique ou polysémique, nous établirons un questionnaire destiné à un panel d'utilisateurs.

Entre d'une part les interactions techniques identifiées en tant que telles (par exemple : connexion, recherche simple ou avancée, forums, etc.) et d'autre part les interactions spécifiquement documentaires (liste des revues, accéder à un numéro, lire une page, etc.), nous chercherons à isoler les points d'interactions qui pourraient conduire à malmener la future interaction documentaire ou mal comprendre la nature de l'interaction technique proposée.

Une analyse des questionnaires permettra de mettre en évidence les interactions ambiguës qui devront/pourront être alors aménagées ou scénarisées.

Analyse des usages par l'analyse des communications asynchrones

La diffusion croissante des outils de communication en réseau, de type synchrone ou asynchrone attachés aux sites portails, s'accompagne de l'émergence de nouvelles pratiques sociales qui méritent une attention particulière. Si les recherches se concentraient hier sur l'analyse du dialogue homme-machine, aujourd'hui elles évoluent de plus en plus vers l'analyse des situations socialement "organisées" en réseaux à travers l'étude des traces laissées par les internautes dans des espaces communication où émergent des situations dans lesquelles circulent des informations, des documents et des nouvelles connaissances au sein d'échanges collectifs.

L'émergence de communautés virtuelles

Dans les premiers temps d'Internet, les sites portails ont d'abord été considérés comme des médias classiques voués à diffuser des informations peu différenciées à des larges audiences. Pendant la croissance anarchique des dotcoms, on a cherché à étendre les concepts du modèle diffusé au cas d'Internet et, par exemple, à raisonner en terme d'audience de sites ou d'exposition des cibles à des messages publicitaires. Toutefois, le web n'apporte pas d'avantages bien clairs sur les journaux, la radio ou la télévision si, justement, on limite son usage à celui d'un média classique. Là, les sites web se montrent peu utiles en se substituant simplement à des services existants. Un avantage majeur de ce "nouveau" média de diffusion (le web) est la possibilité d'inventer et d'offrir des services nouveaux en terme d'interactions, d'échanges informationnels et de communications interpersonnelles autour des thématiques ou des objets présents sur les sites web et ceci dans des points de rencontre centraux de type forums de discussions, chats, blogs, etc. à titre d'exemple, de nombreuses études ont montré que les discussions libres et décousues autour de la machine à café jouent un rôle important dans la circulation de l'information et le renforcement des réseaux interpersonnels tissés dans les entreprises.

L'émergence des communautés virtuelles ou des groupes sociaux autour des espaces de communication synchrones ou asynchrones devient un enjeu stratégique important pour des nombreux sites portails en fournissant différents services permettant à ces communautés de se créer et de fonctionner. Ces services offrent des mises en relations selon des prototypes très variés. Il peut s'agir des listes de diffusion par mël, des forums de discussion, des groupes de discussion (newsgroups par exemple). Tout un ensemble de techniques et de procédures calqués sur les construits d'interactions développés spontanément sur internet.

L'analyse des traces laissées par ces groupements virtuels permettent de faire émerger un certains nombre d'informations pertinentes sur le site le web lui-même : son contenu, son ergonomie et sa conception. C'est justement ces informations qui nous intéressent ici particulièrement. Depuis quelques années nos activités de recherche visent l'analyse des traces des apprenants laissées dans les forums de discussion des plates-formes attachés aux dispositifs de e-learning. Les travaux correspondants nous ont permet de mettre évidence d'une part, les usages des outils attachés à ces plate-forme : informations très utiles pour les concepteurs et les développeurs (Sidir, 2003 ; 2004 ; ; 2005) et d'autres part les comportements des apprenants en termes d'échanges des connaissances dans des environnements partagés : des données utiles aux tuteurs et aux promoteurs du programme e-learning (Sidir et Baudoux, 2005, Sidir, 2006 ; Sidir et Papy 2006)

Méthodologie :

Notre méthodologie sera basée essentiellement sur l'utilisation de deux techniques d'analyse complémentaires :

1. la première technique est empruntée aux analyses des réseaux sociaux (Social Network Analysis) (Scott, 1991 ; Parlebas, 1992 ; Lazega, 1998 ; Lemieux, 1998). Elle permet de mettre en évidence l'existence d'un ou plusieurs regroupements virtuels.
2. la deuxième technique est à l'analyse des traces (contenus) à travers une double lecture (codage) de l'ensemble des messages électroniques déposés par les participants sur le site Persée suivant une grille d'analyse établie en deux étapes :
 1. une première version de la grille sera élaborée à partir de l'étude du modèle de communication en réseau proposé par (Johnson et Johnson, 1998) et d'une recension de la littérature sur l'analyse des interactions (L'Ecuyer 1990 ; Sedlack et al., 1992 ; Huberman et al. 1991 ; Kumpulainen et al. 1999 ; Gunawardena, et al., 1997). Cette étude permettra de mieux rendre compte des divers éléments sociaux, cognitifs et métacognitifs qui constituent une dynamique d'échange dans les regroupements virtuels attachés au site Persée
 2. cette première version de la grille sera confrontée à une première lecture des messages. Cette lecture permettra de relever d'autres caractéristiques afin de compléter la grille d'analyse et d'en supprimer d'autres en évitant de « rentrer de force » toutes les données dans les différentes catégories. Ainsi, nous établirons une deuxième version finalisée et cadrée par le contexte de du site Persée.

Lorsqu'un message portera sur une ou plusieurs catégories, notre pratique consistera à coder chaque partie du message individuellement, ce qui nous permettra de mieux rendre compte des différentes interactions communicationnelles qu'on pourrait trouver dans chaque intervention.

Approche sémiotique par les genres. Structures scénarielles sous-tendant l'organisation du site et impacts sur les types de consultation

Les travaux menés depuis 1996 sur la problématique de la description sémiotique de sites web (et, d'une manière plus large, de documents à support numérique) recouvrent :

- 1/ la mise en place d'une méthodologie de description de sites sous forme de scénarios (de l'organisation du contenu d'un site, de l'expression et de la communication du contenu et de l'appropriation par les utilisateurs (de l'organisation du contenu d'un site, de l'expression et de la communication du contenu et de l'appropriation par les utilisateurs d'un site) ;
- 2/ l'explicitation des modèles culturels (les « standards » ou encore les genres) qui sous-tendent et guident (« motivent ») la production (la conception et la réalisation) de sites et leur appropriation ;
- 3/ la conception et le développement de « modèles de publication » (« templates », en anglais) de sites ou de parties de sites thématiquement ou pragmatiquement circonscrits.

Il s'agit plus particulièrement, d'analyser des corpus de sites web (sites de communes, sites de musées, sites à caractère touristique, sites de recherche, sites de mini-PME, etc.), d'y isoler les principales caractéristiques structurales (tant au niveau du contenu qu'au niveau de l'expression du contenu ainsi que tant du point de vue de la conception que du point de vue de l'appropriation) afin de spécifier des modèles ou scénarios devant faciliter (et aussi standardiser) le processus de la spécification, conception, rédaction et édition d'un site par l'acteur social auquel il appartient (entreprise, musée, labo de recherche, etc.)

Objet et objectifs de l'étude

Dans la description et l'analyse d'un corpus de sites web, il nous paraît très important de distinguer entre :

- 1/ le site comme un lieu de prestation (i.e. la publication et la diffusion du contenu de revues en SHS) ;
- 2/ la prestation elle-même proposée sur ce lieu (i.e. le contenu lui-même) ainsi que les différentes possibilités (services) offertes pour le comprendre, le traiter, l'exploiter, etc.

L'analyse sémiotique du site portail Persée sera appréhendé au travers du deuxième point c'est-à-dire à la prestation elle-même (au contenu des différentes revues actuellement mises en ligne) et aux services (existants et possibles) pour permettre à un utilisateur ou, plutôt à une communauté à s'approprier celle-ci.

Une question centrale qui se pose, à notre avis, est la suivante : tout en conservant au site Persée le caractère d'un site à la fois évidemment patrimonial (du genre « bibliothèque » ou « archives ») et de diffusion « standard » (du genre « infocom »), ne peut-on pas envisager des scénarios pour faire évoluer certaines de ses fonctionnalités vers un site

plus approprié aux exigences et besoin des pratiques de la communication scientifique (didactique, professionnelle, etc.) ?

Cette réflexion met en exergue deux grandes questions théoriques, méthodologiques et, enfin, techniques :

- 1/ peut-on isoler les principaux paramètres structuraux qui interviennent dans le travail de description, d'indexation, d'enrichissement, de traduction, de republication, etc. sur un corpus de documents (textuels, audiovisuels, ...)
- 2/ peut-on spécifier des scénarios de republication (lato sensu) d'un corpus de documents existants sous forme de « nouveaux » documents.

Il s'agit ici d'une problématique certes « compliquée » et « abstraite » mais cruciale pour faire évoluer un site portail vers un vrai lieu au service d'une communauté donnée (dans notre cas : des enseignants, des étudiants, des chercheurs, d'un certain journalisme, etc.).

Méthodologie

- 1/ sélection d'un ou de deux corpus de travail composé d'un certain nombre d'articles scientifiques sur une problématique donnée ;
- 2/ identification et description des principaux sujets et thématiques abordées et mise en place d'une ontologie des domaines sélectionnés ;
- 3/ analyse des genres de discours dans les deux corpus (et, d'une manière stochastique, sur l'ensemble du fond) et spécification d'une ontologie rhétorique et narrative (i.e. d'une ontologie précisant les concepts nécessaires pour construire un discours sur un sujet donné ; cf. 2) ;
- 4/ analyse et description des parties du corpus sélectionnés ;
- 5/ spécification d'un ou de deux modèles de republication ;
- 6/ simulation de processus de traitement (segmentation, description, ... republication) du corpus sélectionné et de sa republication sous forme d'un dossier thématique, d'une hyper-documentation (incluant les ressources audiovisuelles) ou encore d'une documentation multilingue.

Publics ciblés

Pour mener cette étude, un échantillon de 64 "usagers-tests" a été pensé de la manière suivante :

- 16 étudiants en master recherche,
- 16 étudiants en doctorat,
- 16 post-doctorants (ATER),
- 16 enseignants-chercheurs (MCF et/ou PR)

Les usagers-tests seront « recrutés » auprès des universités des membres de l'équipe : Université Paris 8, Université de Lille 3, Université d'Amiens, INALCO/FMSH.

Une étude (toujours) en cours

Les résultats de cette étude sont attendus pour la rentrée 2007 et accompagneront la réalisation de la prochaine version du site.

Bibliographie

- Baccino T., Colombi T. (2004), « Exploration visuelle et navigation dans les hypertextes : quelles stratégies ? », Ergo'IA 2004, Biarritz
- Bruillard E., « Qu'importe qu'ils comprennent puisqu'ils savent s'en servir ! », CNDP, Les dossiers de l'ingénierie éducative sur le web, Les archives de la tribune libre, Avril 2000. http://www.cndp.fr/tice/DossiersIE/texte_ericb.htm
- Chauvin S., (2006) « Les classifications instrumentées par l'hypertexte : améliorer l'exploration dans la structuration des savoirs », ISKO, Semaine de la Connaissance, Nantes 26-30 juin 2006.
- Chauvin S. (2005) « Visualisations heuristiques pour la recherche et l'exploration de données dynamiques : L'art informationnel en tant que révélateur de sens ». Thèse en sciences de l'information et de la communication, Université Paris 8, Décembre 2005, 285 pages.
- Ciaccia, A., Martins, D., « d'informations sur le web : Etude de l'influence de facteurs liés à l'interface, à l'utilisateur et à la tâche » in « Alternatives en sciences cognitives, enjeux et débats », RSTI série RIA, Volume 19, Hermès Lavoisier, 2005
- Corson Y., Aspects psychologiques liés à l'interrogation d'une base de données, Rapport de recherche de l'INRIA – Rocquencourt, 78 pages - Avril 1982, <ftp://ftp.inria.fr/INRIA/publication/publi-pdf/RR/RR-0126.pdf>
- Dinet, J., Rouet, J-F., « La recherche d'information : processus cognitifs, facteurs de difficultés et dimensions de l'expertise », in IHM et recherche d'informations, C. Paganelli (dir), Traité STI, Paris, Hermes, 2002, pp 133-161.
- Dittrich Y., Eriksén S., Hansson C. « PD in the Wild: Evolving Practices of Design in Use ». In: T. Binder, J. Gregory, I. Wagner Proceedings of the Participatory Design Conference 2002, Malmö, Sweden.
- Folcher V., Sander E. (2005) « Usages et appropriation : de l'analyse a priori à l'analyse de l'activité instrumentée », In P. Rabardel & P. Pastré (eds) Modèles du sujet pour la conception, dialectiques activités développements, Octarès, 129-155.
- Folcher V., Rabardel P. (2004) “Artifacts as design-for-use propositions for design-in-use activity”, Invited Conference to the Symposium « Perception, Communication and Activity », International Congress of Psychology (ICP 2004), Pekin, 7-13 Août.
- Henderson A., Kyng M., « There is no place like home - continuing design in use » In J. Greenbaum & M. Kyng (Eds.). Design at Work: Cooperative Design of Computer Systems, Lawrence Erlbaum, Hillsdale, New Jersey, 1991, pp. 219-240.
- Jannès-Ober, E., «L'utilisateur face à la bibliothèque numérique : l'expérience du portail d'information scientifique de l'Institut Pasteur », Les Bibliothèques Numériques, Hermes, 2005, pp. 51-64
- Jeanneret, Y., « Y a-t-il (vraiment) des technologies de l'information ? », Presses universitaires du Septentrion, 2000)

- Le Marec J., « Dialogue ou Labyrinthe : la consultation des catalogues informatisés par les usagers », Etudes et Recherches, BPI Centre George Pompidou, 1990.
- Lompré N. (1999) « Usages de conception des pages Web et recommandations ergonomiques », ICUST 99, Usages et services des Télécommunications, Bordeaux 7-9 Juin 99.
- Lucas N., Sidir M., Giguët E. (2006) « Analyse discursive des forums pédagogiques » In E., Bruillard, M., Sidir et M. G-L., Baron (Coord.), Communications et interactions dans les forums de discussion éducatifs, 2006, Editions Octares, Toulouse
- Papy F., Chauvin, S. (2006) « Une proposition pour améliorer la performance de l'utilisateur au sein du système d'information global de la bibliothèque », ERGO-IA 2006, 11-13 octobre 2006, Biarritz
- Papy F. (2006) « Modélisation des meta-données bibliographiques : une contribution instrumentée destinée à améliorer la circulation des usagers dans les connaissances de la bibliothèque », IC'2006, 28-30 juin 2006
- Papy F., Chauvin S. (2005) « Au-delà de la transfiguration du catalogue. Le Visual...Catalog », BBF 2005, Paris, t. 50, n° 4, p. 5-12
- Papy F., Chauvin S. (2005) « Pour une approche visuelle et ergonomique dans la recherche et l'exploration d'informations au sein d'un OPAC de SCD. L'exemple du Visual...Catalog », in Bibliothèques Numériques, Traité IC2, série Management et gestion des STICS, Hermès Lavoisier, 2005
- Papy F., Folcher V., Sidir M., Cerrato-Pargman T. (2004) « E-learning et technologies pour la coopération : inadéquations artefactuelles et logiques des activités instrumentées », Communication au colloque Ergonomie et Informatique avancée, ERGO IA', 17-19 Novembre 04, Biarritz.
- Pavé F., L'illusion informaticienne, L'Harmattan, 1989
- Sidir M., Mariette H. (2006) « L'apprentissage collaboratif : le temps de l'évaluation », dans TICE Méditerranée, Gênes (Italie), 26 et 27 mai 2006
- Sidir M., Boudaux S. (2005) « La médiation par les TIC de la communication éducative », H2PTM'05, Hermès-Lavoisier, pp. 395-408
- Sidir M., Cochard G-M. (2004) « Méthode et outils de gestion en e-formation ». In Enseignement à distance : épistémologie et usages, Hermès-Lavoisier, p. 92-114
- Sidir M. (2004) « Modes de collaborations au sein de groupes d'apprentissage dans une formation à distance universitaire », TICE 2004, Compiègne, 20-22 octobre 2004
- Simonnot, B., "De la pertinence à l'utilité en recherche d'information : le cas du Web", actes du colloque "Recherches récentes en Sciences de l'information", 21-22 mars 2002, Toulouse, Adbs-Editions, V. Couzinet et G. Régimbeau (dir), pp 393- 412
- Stockinger P. (2005) « Les sites web. Description, évaluation et production », Science Publications Paris, Hermes Science Publications, 240 pages

- Stockinger P. (2004) « Le document audiovisuel. Description et exploitations pratiques », Paris, Hermes Science Publications, 240 pages
- Stockinger P. (2003) (eds.) « Portails et collaboratoires pour l'enseignement et la recherche », Paris, Hermes Science Publications, 230 pages
- Stockinger P. (2001) « Traitement et contrôle de l'information. Perspectives sémiotiques et textuelles », Paris, Hermes Science Publications (195 p)
- Stockinger P., Nikitenko C., (2000) (eds.) « La publication en ligne », Paris, Hermes Science Publications, 225 pages
- Tricot, A., Tricot, M., « Un cadre formel pour interpréter les liens entre utilisabilité et utilité des systèmes d'information ». Colloque Ergo-IHM 2000, Biarritz, 3-6 octobre.