

L'enjeu des métadonnées dans un contexte de "redocumentarisation"

Evelyne Broudoux, Claire Scopsi

► **To cite this version:**

Evelyne Broudoux, Claire Scopsi. L'enjeu des métadonnées dans un contexte de "redocumentarisation". Études de communication - Langages, information, médiations, Université Lille-3, 2011, pp.9-22.
<sic_00713491>

HAL Id: sic_00713491

https://archivesic.ccsd.cnrs.fr/sic_00713491

Submitted on 1 Jul 2012

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.

Métadonnées sur le web : les enjeux autour des techniques d'enrichissement des contenus

Evelyne Broudoux

EA 4420, DICEN

(Dispositifs d'Information et de Communication à l'Ere Numérique)

Université de Versailles Saint-Quentin – IUT de Vélizy

Claire Scopsi

EA 4420, DICEN

(Dispositifs d'Information et de Communication à l'Ere Numérique)

Conservatoire national des Arts et Métiers - Paris

1. L'enjeu des métadonnées dans un contexte de « redocumentarisation »

En 2006, l'auteur collectif Roger T. Pédaque traduisait par le terme « redocumentarisation » le phénomène d'une transformation des fonctions documentaires par le processus de numérisation provoquant la révision d'un ordre documentaire antérieur et l'hypothèse d'une entrée dans une « seconde modernité » (Beck, 2001) ; la « première modernité » étant le résultat du développement de l'imprimé. Cette reconfiguration comporte des enjeux considérables puisqu'elle ne touche « rien de moins en effet, que les artefacts qui nous permettent d'affirmer notre identité, de participer à une communauté, de régler nos transactions ou encore de maintenir notre mémoire sociale » (Pédaque, 2006, 10).

Nous rappellerons dans un premier temps les concepts d'analyse de la redocumentarisation que nous considérons pertinents pour aborder la production actuelle de métadonnées documentaires tout en nous appuyant sur quelques définitions issues de la littérature récente (métadonnées, indexation sociale, folksonomies) afin d'évoquer les enjeux de cette activité pour l'accès aux ressources informationnelles du web.

Dans un second temps nous présenterons les contributions réunies dans ce dossier en les reliant aux concepts de la redocumentarisation. Enfin, nous dresserons un bilan de ce dossier en proposant des pistes d'approfondissement.

1.1 l'hypothèse d'une « redocumentarisation du monde »

De l'hypothèse formulée par Pédaque d'une redocumentarisation nous retiendrons tout d'abord la recomposition des médiations documentaires, par le biais notamment de la télédocumentation qui donne l'initiative au lecteur, jusqu'à donner l'impression d'un effacement des médiations. La malléabilité du document numérique, la création de multiples outils permettant aux usagers d'agir sur le document, conduisent à un rééquilibrage des rôles d'auteur-émetteur et lecteur-récepteur, tout en modifiant notre perception de la notion de document original. Ce phénomène rend plus crucial la communication, en marge du contenu lui-même, d'éléments décrivant pourquoi, comment et par qui les états successifs de ce contenu ont été élaborés.

Les transformations des dispositifs de productions et d'échanges documentaires induisent des mutations dans la chaîne des médiations humaines. Les processus dynamiques, mobilisant un grand nombre d'acteurs, construisent une « intermédiatisation ». S'y trouvent modifiés les rôles investis, les pouvoirs et statuts attribués aux individus et institutions producteurs ou propagateurs de documents ainsi que leurs modes de régulations, économie d'autant plus complexe que les anciens et nouveaux médias coexistent et s'interpénètrent le plus souvent dans la chaîne de production d'un même document.

Ces mutations posent la question des savoirs et savoir-faire nécessaires pour produire des documents ou y accéder, mais aussi des compétences critiques permettant d'appréhender les contraintes ou les conventions, variables selon les processus éditoriaux, qui affectent le discours.

Nous intéressent également la transformation sociale qui se manifeste par de nouvelles superpositions entre sphères privée, collective et publique et le renforcement de deux pôles, celui de l'autonomie individuelle (publicité de soi) et celui du contrôle social (panoptique réticulaire).

1.2 Quelques définitions associées aux données

Nous retenons pour ce dossier que les possibilités de classement, de mise en relation sémantique, de découpage, recherche dans les textes et les images, les annotations partagées changent la relation au document et que les quatre propriétés du document (mémorisation, organisation, création et transmission) s'en trouvent démultipliées et transformées. Sous cet angle, nous proposons d'explorer le rôle joué par les métadonnées, c'est à dire une composante particulière du document dans ce processus de transformation.

La maîtrise de l'organisation des contenus et des systèmes de représentation de l'information est liée à la gestion des métadonnées. Ces objets conceptuels articulant documents, connaissances, services, langages contribuent aux enjeux d'innovation pour une large gamme d'acteurs impliqués dans l'éducation, les industries culturelles, l'édition, la presse et le développement logiciel.

Selon Anne J. Gilliland (2000) le terme metadata, littéralement les données sur (ou à propos) des données, est complexe car compris dans des acceptions diverses par différentes communautés professionnelles : dans le monde des bibliothèques ou des archives il désigne largement toutes les informations applicables à un document, qu'elles soient produites manuellement par des professionnels et réunies dans un document secondaire (une notice bibliographique par exemple) ou isolées automatiquement par extraction de concepts. Ces deux méthodes présentent cependant des différences notables puisque la production manuelle permet une indexation interprétative (Amar, 2000) c'est-à-dire d'ajouter des informations « extrinsèques » au document, alors que l'extraction ne permet que d'isoler des métadonnées intrinsèques (par exemple des termes inclus dans le contenu) ou déductibles du contenu (des calculs d'occurrences ou de co-occurrences, extraction d'entités nommées) (Monnin, 2009).

Dans les communautés techniques le terme métadonnées désigne des données opérationnelles produites automatiquement par les systèmes informatiques : désignations ou emplacement de fichiers, informations structurelles (tailles, formats) données de repérage temporel (horodatages) ou géographiques (données issues des SIG).

La normalisation des formats (ou schémas) de métadonnées, qu'elles soient produites manuellement ou automatiquement, joue un rôle majeur dans l'organisation de l'accès aux documents ou aux ressources informationnelles (parties de documents) : les formats MARC ont conduit au partage des métadonnées bibliographiques et modifié les pratiques quotidiennes des bibliothécaires, en permettant le développement des catalogues collectifs.

Pour les documents audiovisuels la normalisation des formats de métadonnées d'images, encapsulées dans un fichier « conteneur » associé au fichier multimédia est un verrou économique partiellement levé. Les formats de description IPTC et EXIF permettent, lors du traitement documentaire des photographies, la récupération automatique et fiable des informations produites par l'appareil. Dans le domaine de l'image animée des formats fédérateurs permettraient, en transférant les métadonnées de production avec les vidéos, de créer plus facilement des services de diffusion à valeur ajoutée. Malheureusement cette

standardisation tarde à émerger et bloque le développement d'outils intégrant toutes les activités de la chaîne audiovisuelle. (Rault, 2010)

La question de l'accès partagé aux ressources informationnelles du web, s'appuie aussi sur des schémas standards de métadonnées, facilitant la « collecte » et l'interprétation par les moteurs de recherche des descriptions inscrites dans les balises <META> du fichier HTML. Le schéma Dublin Core (norme internationale ISO 15836) élaboré depuis 1995 a ainsi permis la mise en commun de ressources internationales et est utilisé notamment dans les projets d'archives ouvertes OAI, cependant les principaux moteurs du web ne le prennent pas en compte et son utilisation reste confinée aux communautés scientifiques.

Deux axes de réflexion ont émergé ces dernières années pour produire des points d'accès aux informations de plus en plus volumineuses publiées sur le web : l'exploitation des traces et l'indexation collaborative. Par « exploitation de traces » nous désignons l'activité consistant à récupérer les informations générées par les individus lors de leur activité sur le web. Ces traces peuvent être involontaires voire inconscientes (requêtes d'interrogation, traces de navigation etc.) ou produites volontairement dans le cadre d'une activité spécifique, mais sans conscience de leur réutilisation possible (contenus des blogs, ou des réseaux sociaux, commandes sur les sites marchands...). Ces traces d'activités sont exploitées à l'instar des métadonnées (mais peut-on alors parler de métadonnées ?) pour produire de nouvelles formes d'accès à l'information. La prise en compte par le moteur de recherche Google des liens pointant vers une page dans le calcul du « PageRank » d'une page web est un des premiers exemples de ce type de détournement. Les suggestions d'achat selon le profil de l'utilisateur calculé d'après ses requêtes précédentes en sont un autre.

Par « indexation collaborative », nous entendons les dispositifs permettant aux usagers d'affecter des mots-clés à des contenus pendant leurs cycles de publication et consultation web. L'auto-indexation qui consiste à apposer un tag à un contenu pour le caractériser devient un acte de publication pris en charge par lecteurs et auteurs, encourageant une attitude d'appropriation du web et de ses méthodes de référencement (Broudoux, 2005).

La production croissante de « Users Generated Content » (UGC) ou données produites par l'utilisateur, dans des sites aux caractéristiques assez diverses grossièrement regroupées sous le terme de web 2.0 (blogs, sites de partage de ressources, applications collaboratives...) produit une forme particulière de métadonnées exploitables pour l'accès à l'information (Ertzscheid et Gallezot, 2006) : les tags appliqués par les utilisateurs soit dans une logique « communautaire » (l'utilisateur fait don de ses métadonnées à une communauté pour construire un système de recherche performant), soit dans une logique personnelle (l'utilisateur applique des tags sur une ressource pour la retrouver plus commodément) soit dans une logique promotionnelle (pour inciter un public nombreux à accéder à la ressource). L'objectif de partage de ressources s'étant élargi récemment au partage de l'expérience immédiate avec Twitter.

Les folksonomies, littéralement des « taxonomies sociales », résultent de l'agrégation des tags ainsi produits pour construire des « nuages », cartographies, classifications ou systèmes de filtrage sémantique (Zacklad, 2007).

C'est cet ensemble de productions englobées sous le vocable « métadonnées » et porteuses de valeur ajoutée pour le web que nous proposons de confronter au paradigme de la redocumentarisation.

2 Présentation du dossier

2.1 Les intentions : repérer les nouveaux modes d'accès aux contenus orchestrés par les métadonnées

L'objectif général de ce dossier était de repérer les nouveaux systèmes d'organisation et de représentation des contenus sur le web ainsi que les modes d'accès à l'information en voie de diversification, orchestrés par les métadonnées. Étaient ainsi concernées les différentes techniques d'adjonction, de filtrage, de classification, de tri, d'infographie, de cartographie.

Quatre axes de réflexion étaient proposés :

- un axe social, s'intéressant aux acteurs de la production des métadonnées, aux nouveaux rôles endossés par les usagers dans les pratiques d'indexation sociales ; aux questions soulevées par l'exploitation des identités numériques,
- un axe organisationnel consacré aux travaux de régulation et normalisation facilitant l'interopérabilité de ces données,
- un axe économique s'interrogeant sur les enjeux commerciaux liés à la production et l'exploitation des métadonnées et leur place dans l'économie du numérique,
- un axe « pratiques » rendant compte des usages observables.

2.2 Les contributions :

L'axe social et l'axe des pratiques se sont révélés fructueux avec la contribution d'**Olivier Le Deuff**, et celle de **Nathalie Casemajor**. Le premier aborde la question de la recomposition sociale et de l'acquisition des compétences critiques, en établissant le constat d'une culture nouvelle de la folksonomie, liée au nouveau rôle de l'utilisateur du web 2.0 placé au centre d'un processus d'organisation et de catégorisation des ressources du web apparemment simplifié, mais qui comporte des complexités implicites. L'auteur conclut à la nécessité d'apprentissage d'une « information literacy », pour acquérir une maîtrise dans la gestion des métadonnées de soi et éviter leur captation non désirée.

La seconde interroge la figure de l'utilisateur actif, dans la lignée de Certeau, en analysant les dispositifs collaboratifs proposés autour des fonds photographiques de Bibliothèque et Archives Canada. Elle relativise le paradigme de l'efficacité documentaire en reconnaissant à la fonction de « documentarisation collaborative » un rôle de compensation sociale lorsque les usagers issus des minorités culturelles sont invités à se réapproprier leur histoire en contribuant au taguage des documents les concernant. Dans cet exemple de recomposition des médiations, les phénomènes d'intermédiatisation qui concernent les processus d'échanges et de production des documents sont aussi importants que les documents eux-mêmes.

Enfin, **Fabrice Pirolli** recommande l'indexation collaborative comme outil de veille où les métadonnées considérées comme intéressantes deviennent des moyens de détection des signaux faibles lors de la phase d'identification des axes de surveillance d'un processus de veille. À partir du calcul de co-occurrence relative entre deux tags, est modélisé un outil de veille basé sur le suivi dans le temps de co-occurrences sélectionnées dans un ensemble d'utilisateurs partageant des ressources jugées pertinentes.

Ces exemples sont caractéristiques de l'importance prise par le contexte dans la production de métadonnées liées à la caractérisation des contenus. Dans les deux articles suivants, ce sont les métadonnées sémantiques qui sont appelées à jouer un rôle déterminant dans les manipulations liées à la fonction mémorielle et organisationnelle des documents.

Ainsi pour **Vincent Ribaud, Philippe Saliou et Claude Laporte** qui présentent les techniques mises en œuvre dans la création d'un assistant de mémoire pour les très petits projets d'ingénierie logicielle (tpo), les difficultés dans la gestion des métadonnées documentaires tiennent à leurs relations complexifiées. L'utilisation combinée d'un wiki sémantique qui génère ses propres métadonnées et d'une ontologie de référence dans le domaine culturel servent à la conception d'un outil de gestion de projet où l'utilisateur annote et précise les liens unissant une tâche et un artefact. Pour que tiennent ensemble dans une chronologie temporelle, les différentes versions d'un même document, ses différents formats

et ses différentes traductions, le projet différencie le document comme œuvre, comme expression, comme manifestation ou comme item, en s'appuyant sur une distinction établie par l'Ifla (*International Federation of Library associations and institutions*), ce qui permettra la gestion des différentes « incarnations » d'un document.

Dans un autre registre, **Nathalie Pinède et David Reymond** proposent une méthode d'analyse des pages d'accueil des sites web d'un même domaine en étudiant la sémantique de leurs liens internes. L'objectif est de répertorier des lexiques susceptibles d'influencer ultérieurement le référencement sémantique de sites web. La métadonnée conceptualisée ici est l'unité lexicale hypertexte (UHL), c'est-à-dire la composante de base des menus d'aide à la navigation élaborés et catégorisés par les éditeurs de sites. La proposition est doublée d'un outil collaboratif permettant d'enrichir la taxonomie produite. Les représentations synthétiques des contenus générés à partir des choix terminologiques antérieurs servent à ajuster les stratégies éditoriales des organisations concernées.

3 Quelques pistes de réflexion

Les contributions de ce dossier nous invitent à prolonger la réflexion dans plusieurs directions, interrogeant les propriétés de mémorisation et d'organisation des idées, ainsi que de créativité et de transmissibilité qui sont amplifiées par les évolutions que subit le document à l'heure du numérique. Seraient à interroger :

- **La temporalité de la documentarisation** : la distinction entre les métadonnées produites avant la publication, et celles produites à partir des documents déjà publiés tend à s'estomper : Pinède et Reymond proposent un dispositif hybride dans lequel les lexiques construits a posteriori sont un support à l'élaboration collaborative d'une taxonomie, par un groupe identifié de contributeurs.

- **L'objet même de la documentarisation** : que redocumentarise-t-on ? s'interroge Le Deuff. S'agit-il de mesurer ou optimiser l'efficacité des tags pour la description de ce qu'ils sont supposés décrire ou de produire des descriptions parfois involontaires de nouveaux objets. En même temps que se sont répandues les techniques de tagging en ligne, l'usage des métadonnées s'est vulgarisé et a quitté le domaine des professionnels de la documentation et de l'organisation des connaissances pour s'étendre aux techniques de référencement web, de gestion de contenus, de commerce en ligne et plus largement aux actes d'enregistrement des traces de lecture par les internautes eux-mêmes pendant leur consultation de sites web. Cette situation auto-médiatrice entraîne des formes de documentation de soi non maîtrisées par l'internaute. La transformation des traces « pour mémoire » (hypomnemata) que constituent les annotations diverses (mots-clés, commentaires, etc.), captures d'écran ou surlignages accompagnant la lecture sur le web en « legenda » « récits sur soi » résultant de l'addition des tags et traces produits par un individu, constitue une conséquence non maîtrisée des usages du web et leur utilisation par certaines entreprises pour construire et exploiter des profils est un détournement de leurs fonctions initiales.

- **La gradation à établir de l'intelligence collaborative à l'intelligence sociale**. En effet, de l'une à l'autre, il ne s'agit pas seulement d'un changement d'échelle, mais aussi du degré de conscience des acteurs. Si les « User generated Content » relèvent bien d'une intelligence collective, on ne peut cependant parler de collaboration que dans la mesure où les acteurs ont conscience d'œuvrer pour un projet commun et unissent et coordonnent leurs efforts pour l'accomplir, situation qu'il convient de distinguer de la notion de « filtrage collaboratif », populaire sur le web, qui réfère en réalité à des algorithmes de détection de voisinage pour la recommandation servant directement à l'exploitation, a posteriori, et indépendamment de la volonté des usagers, des traces laissées par leurs activités sur le web. Les pratiques évoquées par Ribaud (co-construction d'une mémoire de projet), Casemajor (ajout d'informations sur

une collection photographique) qui se caractérisent par une intention collective, sont donc à distinguer des exploitations décrites ou expérimentées par Le Deuff (description intentionnelle de sites, induisant une description non intentionnelle des auteurs des tags), Piroli (utilisation des folksonomies intégrées dans les applications de signets, pour identifier des objets de veille), et, partiellement, Pinede et Reymond (production par les éditeurs de sites web d'outils de navigation – des menus hypertextes - que l'on utilise a posteriori pour élaborer une taxonomie des sites).

- **Le caractère illusoire, d'une entière externalisation du processus de redocumentarisation.** Plusieurs contributions marquent une étape intéressante dans la considération de l'autonomie des productions collaboratives : la prise de conscience du rôle de l'animation pour intensifier, coordonner, motiver la collaboration. Pinede et Reymond, Ribaud et alii, Casemajor, décrivent des dispositifs en semi-autonomie, où le processus individuel se prolonge dans une réflexion collective coordonnée. Casemajor explique le « relatif échec de l'initiative BAC » par le fait que « la stratégie d'animation et l'investissement de moyens humains nécessaires à ce type de projet aient été négligés ou sous-estimés ». La délégation complète de la coordination aux outils informatisés ou à des contributions bénévoles totalement autonomes, si séduisante économiquement, fait place à un encadrement éditorial des contributions sous Flickr et à une stratégie d'animation avec un accompagnement professionnel.

- **La distinction entre efficacité documentaire et statut symbolique de collaborateur.** Le moteur de l'indexation sociale réside dans la recherche d'une efficacité puisée dans les vertus de la collectivité selon des postulats divers :

- a) la force collective est convoquée pour compenser une incapacité matérielle ou économique à mobiliser les professionnels traditionnels de la documentation en nombre suffisant.
- b) l'utilisateur est investi de qualités spécifiques échappant aux professionnels (labilité, liberté, miroir du changement). Le caractère informel et individuel de la contribution sociale lui confère une agilité et une aptitude à épouser les terminologies et préoccupations émergentes, propriété particulièrement adaptée à une recherche de signaux faibles dans un processus de veille (Pirelli).
- c) L'utilisateur contributeur est détenteur d'expertises spécifiques (capacité à identifier des objets, à témoigner à décrire une situation (Casemajor).

Cependant, dans le cas de la redocumentarisation d'un fonds de documents Inuit, Casemajor identifie des situations où l'enjeu de la collaboration ne réside pas seulement dans l'efficacité et l'intensité de l'activité documentaires mais aussi dans l'acte symbolique d'octroyer à l'utilisateur le statut de témoin, d'auteur d'une historicisation ; le dispositif de médiation, qui joue le rôle de compensation ou de réparation identitaire, prime alors sur l'efficacité documentaire.

D'une manière générale, les retours du terrain manquent encore pour décider si les objectifs du marketing des logiciels en ligne directement basé sur la valorisation de métadonnées co-produites ou générées par les « utilisateurs » sont atteints. Les « pratiques » communicationnelles et l'utilisation des métadonnées sur le web devraient faire l'objet de recherches inscrites dans la durée pour en évaluer la portée.

- **Les métadonnées : gestion et normalisation, un enjeu pour les prochaines années?**

La place prise par la production et l'exploitation des métadonnées dans l'économie du numérique est absente de ce dossier. Le marché des métadonnées professionnelles manque au panorama que nous avons voulu établir alors que les acteurs de la gestion de contenus ont depuis longtemps pris place dans le champ éditorial et qu'il aurait été intéressant d'étudier leur rôle dans la gestion et la valorisation des métadonnées liées aux biens informationnels. Remarquons avec des exemples comme « immatériel.fr » ou « crevilles.org » que l'édition

traditionnelle a fini par emboiter le pas de l'édition scientifique et qu'un acteur – l'agrégateur de contenus – a pris une place déterminante dans la gestion des collections et la distribution des biens culturels sur le web (ex : moteurs de nouvelles). Les métadonnées sont devenues déterminantes dans la structuration des échanges transactionnels de l'e-commerce : il suffit pour s'en convaincre d'observer la place prise par les comparateurs de prix. Si la frontière temporelle s'estompe entre métadonnées insérées avant la publication et celles générées après la publication, un penchant se confirme : celui de la création de la valeur après publication sur le web.

La tendance est à la complexité dans les recherches liées à la production de métadonnées documentaires. Sur le web, le champ est riche et les initiatives nombreuses et se situent souvent aux frontières de l'informatique et de la communication humaine dans l'objectif d'assurer un accès pérenne aux ressources. Une interdisciplinarité s'avère indispensable pour comprendre le phénomène de sémantisation du web qui se situe en tension entre processus automatisés de génération de métadonnées structurelles et caractérisation de l'information réalisée par les usagers eux-mêmes.

Pour conclure cette réflexion entamée sur les enjeux liés aux techniques de description et de structuration des contenus et des documents, nous remarquons que l'enchevêtrement des pratiques avec les techniques relatives aux métadonnées rend indispensable de poser un vocabulaire précis sous peine de transformer les métadonnées en un mot-valise transportant un vrac communicationnel. En particulier, les caractéristiques de l'intermédiatisation devraient pouvoir être sériées : les jeux de temporalité, la conscience et l'intentionnalité des acteurs, la couverture des dispositifs et leur ambition, leur caractère manuel ou (semi) automatisé, les objectifs documentaires poursuivis sont des variables qui doivent être définies pour rendre compte de la multiplicité des situations.

Bibliographie

Amar M. (2000), *Les fondements théoriques de l'indexation : une approche linguistique*, Paris : ADBS Éditions, 355 p.

Beck U. (2001), *La société du risque. Sur la voie d'une autre modernité*. Paris, Aubier.

Broudoux E., Grésillaud S. Le Crosnier H. Lux-Pogodalla V. (2005). Construction de l'auteur autour de ses modes d'écriture et de publication in H2PTM'05, pp. 123-142. Hermès/Lavoisier. http://archivesic.ccsd.cnrs.fr/sic_00001552_v1/, date de la dernière visite : 31 janvier 2011.

Calderan L., Hidoine B., Millet J. (2008), *Métadonnées : mutations et perspectives*. Paris, ADBS.

Calhoun K. (2006), *The changing nature of the catalog and its integration with other discovery tools*. Final report (17 mars 2006) prepared for the Library of Congress. Cornell University Library.

Charlet J., Salaün J.-M. (2007), *Comprendre et maîtriser la redocumentarisation*, in La redocumentarisation du monde. Cépadués éditions.

Chartron G., Noyer J.-M. (1999), *Normes et documents numériques: quels changements ?* Solaris, n° 6, Décembre 1999 / Janvier 2000, <http://biblio-fr.info.unicaen.fr/bnum/jelec/Solaris/d06/index.html>, date de la dernière visite : 31 janvier 2011.

Chaudiron S., Ihadjadene M., Maredj A. (2008), *La fragmentation et l'unité documentaire en question*, in actes du 16ème Congrès de la SFSIC, 11-13 juin 2008, Compiègne,

http://hal.archives-ouvertes.fr/docs/00/46/87/96/PDF/Chaudiron-Ihadjadene-Maredj_SFSIC-2008.pdf, date de la dernière visite : 31 janvier 2011.

Couzinet V., Chaudiron S. (coord.) (2008) *Organisation des connaissances à l'ère numérique*, Sciences de la société n°75.

Ertzscheid O., Gallezot G. (2006), *Etude exploratoire des pratiques d'indexation sociale comme une renégociation des espaces documentaires. Vers un nouveau big bang documentaire ?* in Chartron G. & Broudoux, E. (Eds). Document numérique et société, ADBS Éditions, <http://archivesic.ccsd.cnrs.fr/>, date de la dernière visite : 31 janvier 2011.

Fortin J., Carloni O., Leclère M., Weiser S. (2009), *Extraction et exploitation de données temporelles pour un portail d'e-tourisme*, in Fouille de Données Temporelles - Analyse de Flux de Données, 9èmes Journées Francophones Extraction et Gestion des Connaissances, 2009. <http://www.mondeca.com/index.php/fr/content/download/444/3399/file/09-%20Fortin-Carloni-Lecl%C3%A8re-Weiser-Extraction%20et%20exploitation%20de%20donn%C3%A9es%20temporelles%20pour%20un%20portail%20d%E2%80%99e-tourisme.pdf>, date de la dernière visite : 31 janvier 2011.

Gilliland-Swetland A. J. (2000), "Setting the Stage: Defining Metadata" in Introduction to Metadata: Pathways to Digital Information, Murtha Baca, Los Angeles: Getty Information Institute.

Gradmann S. (1998), *Cataloguing vs. Metadata: old wine in new bottles?* 64th IFLA General Conference, August 16 - August 21, 1998, <http://archive.ifla.org/IV/ifla64/007-126e.htm>, date de la dernière visite : 31 janvier 2011.

Laborie, S., Manzat, A.-M., Sèdes, F. (2009), *Création et utilisation d'un résumé de métadonnées pour interroger efficacement des collections multimédias distribuées*, 27ème Congrès Informatique des Organisations et Systèmes d'Information et de Décision (INFORSID 2009), Toulouse, <http://hal.archives-ouvertes.fr/docs/00/38/99/05/PDF/INFORSID09-Laborie-Manzat-Sedes.pdf>, date de la dernière visite : 31 janvier 2011.

Mahé A., (2009), *Métadonnées : mutations et perspectives*. BBF, t. 54, n° 1, p. 94-95.

Monnin A. (2009), *Qu'est-ce qu'un tag ? Entre accès et libellés, l'esquisse d'une caractérisation*, in « Connaissance et communautés en ligne », 20es Journées Francophones d'Ingénierie des Connaissances (IC2009), Hammamet, 25-29 mai 2009, http://ic2009.inria.fr/docs/papers/Monnin_IC2009_41.pdf, date de la dernière visite : 31 janvier 2011.

Rault, E. (2010), *La question des métadonnées*. In « Dossier Vidéos en ligne », C. Scopsi (coord.), Documentaliste-Sciences de l'information, vol. 47, n°4, p. 40-41.

R.T. Pedauque (2006), *Document et modernités*, http://archivesic.ccsd.cnrs.fr/sic_00001741/fr/ date de la dernière visite : 31 janvier 2011.

Zacklad M. (2007), *Classification, thésaurus, ontologies, folksonomies : comparaisons du point de vue de la recherche ouverte d'information*. In actes de la Conférence CAIS/ACSI 2007, Montréal, http://www.cais-acsi.ca/search_fr.asp?year=2007, date de la dernière visite : 31 janvier 2011.