

Zacklad M., Bénel A., Bringay S., Barry-Gréboval C., Charlet J., Treins M., Salzano G., Curé O., Détienne F., Barcellini F., Burkhardt J.-M., Sack W., Guibert S., Darses F., Boujut J.-F., Lortal G., Lewkowicz M., Todirascu A., Turner W., Pierrat M.-J., de Saint Leger M., Processus d'annotation dans les documents pour l'action : textualité et médiation de la coopération, In : R. T. Pédaque (Ed.), *La redocumentarisation du monde*, Toulouse : Cépaduès, 2006.

Processus d'annotation dans les documents pour l'action

Résumé

Ce chapitre traite des documents numériques utilisés comme supports à une pratique collective. Ces documents se caractérisent à la fois par une hyper-fragmentation résultant d'une hétérogénéité temporelle (weblog) et auctoriale (wiki) mais aussi par une multiplication des relations explicites entre ces fragments (« web sémantique » au sens large). Le chapitre propose tout d'abord un cadre théorique autour des notions de document pour l'action et d'annotation pour ensuite analyser l'état des recherches dans différents domaines d'application : médecine, conception mécanique, conception de logiciels libres, documentation scientifique.

Mots-clefs : *Travail coopératif assisté par ordinateur, Gestion des connaissances, Etudes d'usage.*

Abstract

This chapter deals with digital documents used as a medium for collective practices. Documents of this kind are hyper-fragmented because they are authored at a different time (weblog) and even by different people (wiki). But these fragments are also massively linked to one another ("semantic web" in a broad sense). In the beginning of this chapter, a theoretical framework is given with the notions of document for action and annotation. Then, a state of the art is given in several domains: medical care, design in mechanics, design of open source software, scholarly documentation.

Keywords: *Computer supported cooperative work, Knowledge management, Use studies.*

1 INTRODUCTION

L'essor du Web dans un contexte Internet ou Intranet a entraîné une multiplication des pratiques collectives centrées sur le document que celles-ci soient de nature strictement professionnelles (conception médiatisée par des plans et des contrats, prise en charge médicale ou médico-sociale par un réseau de santé, par exemple), ou qu'elles soient moins structurées (activité d'associations militantes ou de communautés d'intérêt partageant une passion commune). Toute une série de nouvelles pratiques centrées sur le document sont ainsi en train d'émerger, allant de la mise à disposition de documents dans des petits systèmes de gestion documentaire (amenés à constituer des « micro bibliothèques numériques » partagées par une équipe) à l'utilisation des annotations accompagnant la circulation ou la rédaction collective du document en passant par l'usage des Wiki et Blogs.

Nous nous intéresserons particulièrement ici aux situations dans lesquelles les documents sont des supports à la coordination d'un collectif distribué engagé dans une activité commune finalisée. Dans ces contextes, le document apparaît comme un ensemble de fragments portés par des auteurs divers et dont le contenu final reste largement indéterminé alors même que sa circulation rapide lui fait déjà jouer un rôle majeur d'information, d'aide à la décision et de preuve.

Dans les travaux de l'atelier nous nous sommes en partie appuyés sur le cadre conceptuel des « Documents pour l'Action » (DopA) (Zacklad, 2004) pour rendre compte de la diversité des pratiques coopératives centrées sur le document et du rôle particulier qu'y jouent les annotations. Ces pratiques sont très diverses et vont du domaine de la médecine aux pratiques informationnelles des chercheurs en passant par la coopération autour de plans dans le domaine de la conception. Elles suscitent des projets de conception variés : nouvelles fonctionnalités logicielles pour la gestion collective des annotations, promotion de standards, etc.

2 PROBLÉMATIQUE : DOPA ET ANNOTATIONS

Le point de vue spécifique qui est le nôtre, élargit la perspective documentaire bien que celle-ci se soit déjà, à travers bon nombre de ses auteurs (cf. par exemple : Brown & Duguid, 1995 ; Briet, 1951 ; Buckland, 1997 ; Pédauque, 2003), émancipée d'un rapport trop exclusif aux textes « classiques ». Elle le fait en s'ancrant résolument dans une

Annotation et coopération

perspective communicationnelle et pragmatique, elle-même appuyée sur le cadre théorique des transactions communicationnelles symboliques (Zacklad, 2005a) qui vise à rendre compte des activités coopératives, notamment médiatisées par divers artefacts symboliques, dans une perspective à la fois communicationnelle, cognitive et socio-économique.

Toute analyse de la notion de document nous semble devoir inscrire ceux-ci dans un processus de communication pour partie différé, au sens des processus asynchrones décrits dans le champ du CSCW, entre des producteurs et des récepteurs liés par des intérêts communs. Le document est analysé comme étant l'objet d'une transaction entre des acteurs impliqués dans un processus d'échange visant à la fois des engagements mettant en jeu leur « self » et des connaissances liées à la production d'une « œuvre » au moins pour partie commune (Zacklad, 2005a).

Les transactions communicationnelles sont médiatisées par des productions sémiotiques qui résultent d'une co-production entre réalisateur et bénéficiaire de la transaction. Mais, du fait de la distribution des activités collectives, réalisateurs et bénéficiaires des transactions communicationnelles peuvent ne pas être présents dans le même cadre spatio-temporel. Cela implique de donner une forme pérenne aux objets de la transaction de manière à ce qu'elle puisse être initialisée, interrompue, réactualisée, répétée dans toutes les configurations de présence/absence du bénéficiaire et du réalisateur. La documentarisation est une des stratégies permettant de palier à la distribution des transactions (Zacklad 2004).

La transcription ou l'enregistrement sont des conditions nécessaires mais non suffisantes de la documentarisation. Une prise de note personnelle ou un enregistrement ponctuel visant à faciliter une activité sémiotique dans une situation transactionnelle donnée pourra avoir une utilité locale certaine. Cependant, s'ils ne font pas l'objet d'un investissement documentaire permettant de soutenir les processus de documentarisation, ces supports ne pourront pas être ré-exploités dans des situations transactionnelles différées et distribuées. Le document désignera ainsi pour nous *une production sémiotique transcrite ou enregistrée sur un support pérenne qui est équipée d'attributs spécifiques visant à faciliter les pratiques liées à son exploitation ultérieure dans le cadre de la préservation de transactions communicationnelles distribuées*. Ces attributs doivent permettre au document de circuler à travers l'espace, le temps, les communautés d'interprétation, pour tenter de prolonger les transactions communicationnelles initiées par ses réalisateurs.

2.1 Production sémiotique coopérative

Mais les activités communicationnelles associées à la coopération structurellement ouverte (Zacklad 2003) soutenue par des supports pérennes génèrent des documents très différents des documents qui sont, par exemple, rangés dans des bibliothèques stockant des « œuvres » au sens traditionnel de ce terme. Ces documents, les *documents pour l'action*, sont pourtant de plus en plus indispensables aux activités collectives coopératives et ils correspondent à un nombre croissant d'usages collectifs des TIC dans le contexte d'actions collectives finalisés. Dans certains cas, quand ils regroupent des productions correspondant à des situations transactionnelles distinctes mais fortement interdépendantes, nous parlerons de « dossiers pour l'action ».

Les exemples de DopA sont très divers et correspondent à de nombreux contextes professionnels. Citons, sans être exhaustif, les documents de conception en ingénierie (mécanique, logiciel...), les dossiers patients en médecine, les documents contractuels dans un contexte d'affaires qui passent du stade de proposition commerciale au stade de contrat en bonne et due forme, les dossiers qualité de plus en plus souvent numérisés, les rapports d'étude dans le conseil en management, les forum d'échange dans le domaine, par exemple, du logiciel libre (Gasser & Ripoche, 2003)...

Leurs caractéristiques principales sont les suivantes :

- Inachèvement prolongé : ils possèdent un statut d'inachèvement prolongé pendant la phase active de la production sémiotique collective, phase durant laquelle nous les appellerons des DopA en évolution (vs DopA stabilisés) ;
- Pérennité : ils possèdent une certaine pérennité associée aux engagements des acteurs à l'égard des contenus sémiotiques véhiculés et à la distribution des transactions, d'où des enjeux liés à leur documentarisation et corollairement à leur stockage et à leur indexation ;
- Fragmentation : au moins pendant leur phase évolutive, ils articulent plusieurs fragments plus ou moins liés sémantiquement (notamment des annotations) qui ne peuvent être intégrés mécaniquement ou de manière organique à l'ensemble du document (cf. infra) ;
- Rapport non trivial aux réalisateurs : les différentes parties du DopA sont le plus souvent produites par différents réalisateurs (plurivocité, pluri-textualité...) qui peuvent avoir des statuts variables dans la situation transactionnelle et donc vis-à-vis de la production sémiotique (droits, engagement, responsabilités...) ;

- Rapport argumentatif non trivial aux autres parties du document : chaque partie possède une relation potentiellement complexe aux autres – modalisation, incertitude, relation argumentative vis-à-vis des autres fragments...

Ces caractéristiques des DopA font qu'ils sont à même de justifier de nouveaux principes d'indexation et de classement encore peu pris en compte dans la profession documentaire mais qui répondraient aux besoins de micro-bibliothèques numériques dédiées à des collectifs restreints. Par ailleurs, ils relèvent de technologies littéraires et documentaires très diverses correspondant à des degrés variables de standardisation de la situation transactionnelle qui induit elle-même une standardisation des rôles joués par les participants et du format de leurs productions sémiotiques.

Dans un ordre de structuration décroissant et sans être exhaustif, on pourra ainsi contraster (1) l'usage de formulaires ouverts, (2) le recours à des plans types (par exemple : des modèles systémiques en ingénierie ou le plan conventionnel d'un contrat juridique), (3) des conventions d'écriture correspondant à un style stéréotypé de type descriptif ou argumentatif, par exemple, dans le contexte de la littérature scientifique, (4) les systèmes de question/réponse utilisés dans les forum et enfin (5) les documents narratifs, dans lesquels les efforts de style visent au contraire, à masquer le plan sous-jacent.

2.2 Importance des annotations dans les DopA

La principale difficulté à laquelle sont confrontés les acteurs qui co-produisent un DopA est le manque d'information sur les paramètres du contexte transactionnel associé à la proposition d'un fragment, correspondant lui même à une offre de transaction. Lorsqu'ils interviennent sur le DopA, ils le font en déposant sur le support des fragments libres associés aux différents moments d'une micro-transaction communicationnelle déposée sur un support pérenne.

Dans le contexte des transactions communicationnelles distribuées et recourant à des supports pérennes, les contributions sont médiatisées par des fragments qui seront liés avec plus ou moins de succès à la production sémiotique principale (par ex. le texte principal). Quand le statut des fragments libres n'est pas établi, ils constituent une production sémiotique accessoire. Ces fragments seront progressivement exclus ou inclus du corps du DopA lors des transactions ultérieures impliquant une reprise et une articulation portant à la fois sur l'expression et le contenu sémiotique. Mais si les fragments libres ne sont pas articulés au fur et à

mesure qu'ils sont déposés sur le support, ce travail de reprise est impossible et le DopA ne peut alors soutenir efficacement les transactions émergentes distribuées associées à l'activité coopérative.

2.2.1 Annotation contributive

Le procédé majeur permettant l'articulation d'un fragment libre avec les autres parties du document est la production d'annotations qui permettent d'explicitier la liaison d'un fragment avec la production sémiotique principale, au moment de l'inscription, de l'enregistrement de celui-ci sur le support ou après coup. De même que toute transcription ou enregistrement sur un support pérenne d'une production sémiotique ne suffit pas à constituer un document, le fragment libre ne constitue une annotation qu'à l'issue d'un processus de documentarisation.

Ainsi, nous définirons l'annotation-fragment ou annotation-contributive comme un fragment de production sémiotique documentarisé, c'est-à-dire doté d'attributs spécifiques permettant sa liaison explicite avec les autres composantes du document. L'activité d'annotation, qui ne se confond pas avec le dépôt d'un fragment libre sur un support pérenne, requiert donc un effort spécifique qui permettra la ré-exploitation des fragments dans le cadre d'une activité individuelle ou collective distribuée en restituant les éléments du contexte transactionnel associés à leur production.

L'articulation caractéristique de l'annotation est son ancrage plus ou moins explicite sur une zone du support traduisant une relation elle-même plus ou moins précise avec un élément du contenu sémiotique. Le fait d'indiquer le réalisateur de l'annotation fait également partie du processus de documentarisation de même que toutes les modalités d'articulation explicites traditionnellement utilisées (date, lieu...). Les annotations-contributives peuvent être très laconiques et codifiées, comme les symboles utilisés dans le domaine de la typographie qui sont des instructions de reprise du texte destinées à un tiers ou certaines annotations codifiées utilisées dans le domaine du dessin technique (Boujut, 2003).

2.2.2 Annotation index

Dans certain cas, les fragments ne viennent pas directement contribuer aux processus transactionnels primaires associés au contenu sémiotique principal mais participent de transactions documentaires (Zacklad, 2004) visant à positionner le document dans un système de classification donné. Ces transactions qualifient le support pour permettre son rangement et sa

recherche. Dans ce cas, les annotations-index ou annotations-classificatoires sont elles-mêmes une part intégrante du dispositif de documentarisation ou de re-documentarisation. Elles offrent de nouvelles prises au document en s'intégrant dans une systématique à caractère sémiotique ou formelle qui permet la mise en relation avec d'autres corpus possédant des ressemblances pertinentes selon le point de vue canonique considéré. Elles peuvent porter sur la totalité du document ou exclusivement sur un fragment particulier de celui-ci. Les annotations définies dans le champ du Web Sémantique sont des annotations-index à caractère formel. Les notices utilisées en bibliothéconomie peuvent également être considérées comme des annotations-index portant sur la totalité d'un document.

2.2.3 Annotation attentionnelle

Enfin un troisième type d'annotation peut être identifié. Ce sont des annotations-attentionnelles, qui ne sont pas toujours destinées à avoir une durée de vie importante et qui visent à focaliser l'attention pendant la ré-exploitation d'un support existant (Guibert *et al.*, 2005). Ces annotations plus éphémères, qui polluent cependant certains supports pérennes dans les bibliothèques publiques, ne sont pas fondamentalement différentes des précédentes quant à leur fonction, bien qu'elles soient moins intrinsèquement créatives que les annotations contributives. Elles s'inscrivent surtout souvent dans un empan temporel différent. Surligner une phrase consiste bien à détacher un fragment parce qu'il participe d'une forme d'exploitation associée au processus de lecture ou de délibération collective en cours dans le cas d'un document partagé. Mais la pertinence de l'annotation peut ne pas excéder la durée de la session de travail courante. Son degré de codification sera alors faible et les participants pourraient souhaiter voir les annotations disparaître à l'issue de la session de travail.

Mais ce n'est pas toujours le cas, certaines annotations-attentionnelles pouvant en effet être considérées comme ayant une valeur durable en permettant une mise en relief de certains fragments utiles lors de l'exploitation réitérée et différée du support. L'annotation-attentionnelle rejoint alors l'annotation-index en contribuant à mettre en relief certains fragments de manière à faciliter leur recherche ultérieure et il peut être utile de progresser dans sa codification pour faciliter sa reprise.

3 ETAT DES RECHERCHES DANS DIFFÉRENTS SECTEURS APPLICATIFS

Les recherches en cours présentées dans l'atelier se déroulent dans différents domaines applicatifs qui posent des questions spécifiques dont la prise en compte est indispensable pour amender et prolonger la réflexion théorique. Ces recherches se développent selon deux perspectives complémentaires. Ils s'agit pour une part d'études d'usage des technologies documentaires qui relèvent à des titres divers de la problématique des DopA. Ces nouveaux usages commencent à induire une transformation significative des pratiques professionnelles sans toutefois parvenir encore de manière totalement satisfaisante à suppléer à la répartition spatio-socio-temporelle des transactions. Ces études permettent de préconiser des évolutions fonctionnelles des environnements documentaires dans le domaine des standards techniques, des normes sociales mais aussi dans celui des caractéristiques techniques dont les logiciels de manipulation documentaire doivent être dotés pour réaliser les différentes opérations d'édition, de lecture, de recherche, de navigation.

Parallèlement, d'autres travaux visent à proposer des briques techniques basées sur des hypothèses cognitives, sociales, organisationnelles, liées à l'exploitation des DopA et des annotations associées. Ces solutions sont elles-mêmes basées sur la définition de normes et de standards ou sur la définitions de fonctionnalités inédites prises en charge par des maquettes ou des prototypes. Les recherches que nous avons regroupées dans l'atelier sont loin de couvrir de manière exhaustive l'ensemble des domaines d'application des DopA mais elles présentent néanmoins une diversité de problématiques pertinente (il nous manque, pour ne citer qu'un exemple, des travaux liés à l'utilisation des DopA dans les domaines juridique et administratif). Elles se répartissent de manière assez équilibrée entre études d'usage et propositions de briques techniques avec une légère prédominance des premières. Les recherches présentées par les membres de l'atelier se répartissent dans les domaines suivants :

- Domaine médical, avec des études à la fois générales sur l'environnement et des propositions de briques techniques visant à répondre à certaines problématiques ciblées autour de l'annotation ;
- Domaine de la conception dans le secteur mécanique et dans le secteur des logiciels Open Source avec des études d'usage permettant la préconisation de solutions techniques inédites ;

- Domaine des bibliothèques numériques avec l'analyse de différentes briques techniques en référence à des hypothèses cognitives sur la compréhension et la gestion associée des annotations ;
- Domaine des pratiques informationnelles dans le domaine scientifique où les annotations contribuent à la définition d'un vocabulaire commun qui pourra ensuite constituer des méta-données.

3.1 Domaine de la santé

L'équipe constituée par S. Bringay, C. Gréboval et J. Charlet (Université d'Amiens et INSERM), dans le cadre du projet DocPatient, mène une étude pratique des usages que les professionnels de santé font des annotations lorsqu'ils manipulent le dossier d'un patient. Ils cherchent à construire un outil d'annotation venant compléter un système de gestion des dossiers patients. L'étude du dossier patient papier (Bringay *et al.*, 2004) leur a permis de délimiter les spécificités fonctionnelles des annotations numériques permettant de reproduire les pratiques papier. Dans un tel dossier, l'annotation est un fragment rédigé par un annotateur (différent ou non du ou des auteurs du document annoté) dont l'intertextualité (distance entre l'annotation et le texte du document annoté) est très variable ;

La description des spécificités fonctionnelles des annotations (visant à reproduire les pratiques papier) ainsi que la description des contraintes juridiques ont permis à l'industriel associé au projet DocPatient de développer un premier module permettant de produire et de consulter des annotations numériques. Celui-ci a par la suite décidé d'exploiter les caractéristiques du support numérique pour proposer de nouvelles pratiques aux professionnels de santé basées sur l'exploitation de ces annotations. Pour cela, les traitements possibles ont été listés :

- création de nouveaux documents éditables en combinant des annotations choisies par l'utilisateur (ex. : pour construire le compte rendu d'hospitalisation),
- création de nouveaux documents (lecture seule) pour naviguer dans le dossier, en combinant des annotations choisies par l'utilisateur,
- filtrage automatique des annotations en fonction des droits d'accès fixés par l'annotateur,
- filtrage manuel des annotations choisies par l'utilisateur,
- envoi à des destinataires des annotations en tant que messages attachés à des documents.

Rapport PSI Pédagogue

Pour appliquer des traitements aux annotations, un ensemble de propriétés leur a été affecté. Ces propriétés peuvent être vues comme des annotations formelles s'appliquant à des annotations informelles. Des tests sont actuellement menés sur le site pilote (juin 2005) (Bringay *et al.*, 2005a ; Bringay *et al.*, 2005b) pour étudier l'impact de ces fonctionnalités sur les pratiques et sur le contenu du dossier patient informatisé.

De manière complémentaire, l'équipe constituée par M. Treins, G. Salzano et O. Curé (Université de Marne la Vallée) travaille également dans le contexte médical et a commencé par une étude visant à mettre en évidence l'ensemble des contraintes annotationnelles portant sur ce type de document :

- les spécificités de ce type de document : comptes rendus, courriers médicaux, prescriptions, résultats d'examens, épisodes de soins, formulaires administratifs, etc.
- contraintes légales : identification et authentification, sécurité et confidentialité des données, gestion des droits d'accès, etc.
- contraintes organisationnelles et techniques : accès à des terminologies et ontologies partagées, intégration d'applications hétérogènes dans un environnement distribué, etc.
- contraintes particulières sur les annotations : selon la nature de l'épisode de soins et de la discipline médicale concernée, cycle de vie des annotations, etc.

A partir de ces contraintes, l'objectif est d'étudier l'adéquation entre des initiatives du secteur de la Santé (HL7, EHRCOM, DICOM, etc.), des ontologies et terminologies du domaine médical (UMLS, SNOMED, OpenGalen, etc..) et de stockage (modèles de base de données). Cette étude doit permettre de mettre en évidence des limites des initiatives existantes et préconiser des axes de solutions effectives et opérationnelles.

3.2 Domaine de la conception

L'équipe constituée de F. Détienne et F. Barcellini, J. M. Burkhardt et W. Sack (INRIA, Université Paris 5, Université de Berkeley) a étudié les discussions en ligne dans des communautés de pratique (communautés OSS) et les pratiques de citation vue comme de l'annotation. Jusqu'à présent, le modèle dominant pour traiter et représenter les discussions en ligne est fondé sur le lien de « réponse à » : les messages sont organisés en fils de discussion. Si ce modèle permet de bien appréhender les

notions de centralité/périphérie des participants dans ces communautés, il est plus faible pour reconstruire la cohérence thématique des conversations.

Barcellini *et al.* (2005) ont montré qu'une autre approche, fondée sur les pratiques de citation électronique, permet d'organiser les discussions en graphes en préservant la cohérence thématique des discussions dans une projet de conception de logiciel OSS. Dans un message électronique, l'alternance de citations et de commentaires ou annotations à ces citations pourrait ainsi être rapprochée de la succession des tours de parole dans une discussion synchrone. En général, un tour de parole a un lien très fort avec un tour de parole précédent (principes pragmatiques de contiguïté et de pertinence) ; par exemple, quand un participant pose une question, quelqu'un va y répondre dans le tour de parole suivant. La citation électronique maintient cette illusion de proximité en incorporant deux tours de parole à l'intérieur d'un unique message.

Deux perspectives de recherche sur la citation électronique peuvent être envisagées :

- la conception d'éditeurs de conversation basées sur l'annotation : on peut alors, en plus de la citation liées au discours, offrir des moyens d'ancrage sur des documents en cours de co-élaboration.
- La construction de mémoire de projet de conception basée sur une analyse des citations dans les discussions en ligne : on peut étendre l'analyse des citations électroniques par l'analyse des liens sémantiques entre citation et annotation (commentaire).

Toujours dans le domaine de la conception, mais cette fois mécanique, une étude psychologique et ergonomique a été réalisée, dans une situation de conception de produit utilisant des outils numériques contemporains, par l'équipe constituée de S. Guibert, F. Darses et J.-F. Boujut, respectivement au CNAM (pour les deux premières) et à l'INPG-GILCO (Guibert *et al.*, 2005). Cette recherche menée dans le cadre de réunions de conception en présentiel avait pour but de définir l'utilisation et la production des annotations en conception de produit. Cette étude a permis de mettre en avant le rôle mineur joué par les annotations textuelles et le rôle majeur des annotations de type graphique. Les annotations textuelles supposées suppléer la déficience des échanges oraux à laquelle conduit la médiatisation des activités (Blanco & Gardoni, 2001) ne suffisent pas à prendre en charge toute la complexité de l'activité. Cette étude a également fait apparaître la variété des annotations graphiques et leur caractère figuratif. Les fonctionnalités d'annotation des outils de CAO actuels (réduites à des annotations de

Rapport PSI Pédauque

type déictique) doivent donc être enrichies. Enfin, il a été mis en avant que la fonction des annotations n'est pas de développer des parties de la solution mais de fournir aux membres de l'équipe des descriptions contextuelles du problème et de la solution. On préconise que les annotations ne soient pas être intégrées au modèle numérique mais conservées sur des calques à part.

Cette étude mériterait d'être approfondie afin d'examiner l'effet de la fonction des concepteurs et de l'organisation sociale du collectif sur la production d'annotations. En effet, il semblerait que, plus les questions à la charge d'un concepteur sont discutées, plus celui-ci produit des annotations. Ceci confirmerait que la fonction des concepteurs est une information importante à conserver dans l'indexation des annotations (Zacklad *et al.*, 2003).

En visant notamment des applications dans le domaine de la mécanique, l'équipe constituée par G. Lortal, M. Lewkowicz, A. Todirascu (respectivement à l'UTT pour les deux premières et à l'Université de Strasbourg pour la troisième) développe un cadre conceptuel et un prototype dans lequel l'utilisation d'annotations et de commentaires glossiques participe à l'interprétation collective et la création collective de sens (Weick, 1979). A l'intérieur des DOPA, l'annotation enrichit le document à plusieurs niveaux. Au niveau computationnel, elle sert au traitement informatique du contenu (ajout de balises pour une structuration du document répondant aux objectifs du Web sémantique, pour la recherche d'information, pour la mise en forme de documents, pour une interopérabilité des services). Au niveau cognitif, l'annotation sert au traitement des informations par l'humain. Le traitement cognitif du contenu peut être aussi bien individuel que collectif.

Pour soutenir les échanges et la construction d'une interprétation collective autour de ce document, cette équipe a entrepris de développer un collecticiel support à une activité de lecture-écriture de type herméneutique. En effet l'annotation est un élément classique de l'herméneutique, un fragment textuel ancré au document qui a donné naissance à une idée critique. Afin d'assister l'utilisateur dans son activité d'annotation, l'équipe adapte un modèle de rhétorique médiévale représentant la création discursive (Carruthers, 1990) à leur problématique de création de fragment textuel (Lortal *et al.*, 2005). Dans le cadre d'une conception itérative du collecticiel, une première version a été développée en se basant sur un serveur (Zannot) respectant le protocole d'annotation du W3C (Annotea).

3.3 Domaine des bibliothèques numériques

L'équipe coordonnée par A. Bénéel, transverse à plusieurs ateliers, travaille dans le domaine des bibliothèques numériques. L'un des enjeux de l'annotation numérique est de redonner une matérialité à ces bibliothèques. Cette orientation pourrait surprendre. En effet, les bibliothèques numériques ont d'abord été vantées pour leur « virtualité » : l'accès au livre « dématérialisé » est quasi-immédiat, mondial et en autant d'exemplaires que nécessaire. Pourtant, la lecture dans ces bibliothèques numériques semble difficile, comme si la compréhension d'un ouvrage nécessitait une « com-préhension » (prise avec soi) du volume matériel.

Quels sont les gestes du lecteur qui participent de cette appropriation du livre ? Parmi les différentes typologies proposées dans la littérature (Mazhoud *et al.*, 1995 ; Veron, 1998 ; Lortal *et al.*, 2005), nous retiendrons en particulier la dernière qui distingue :

- la sélection dans une source documentaire de fragments,
- le regroupement de fragments documentaires dans des collections,
- l'organisation de différentes collections sous forme d'un plan,
- la rédaction d'un brouillon à partir d'un plan,
- la soumission d'une copie à un supérieur ou à des pairs,
- la publication de la version finale.

Ainsi définie, l'annotation s'écarte de la simple prise note et devient « lecture active » : une lecture qui passe par l'écriture, à la fois comme moyen et comme but.

L'instrumentation de cette activité d'annotation au sens large ressemblerait moins aux fonctions offertes par les outils bureautiques courants qu'à celles de systèmes hypermédias tels que *Hyper-G* (Maurer, 1996), *PASTEL* (Tanguy, 1998), *SESAME/RECIS* (Prié, 1999 ; Egyed-Zsigmond, 2003), *TheBrain®* (Personal Brain, 2001), *Porphyry* (Bénéel, 2003) et *LUCIA* (Perlerin, 2004).

Des membres de l'équipe analysent actuellement ce petit panel d'outils afin de voir comment chacune des actions recensées est rendue possible par les modèles de données et les interfaces homme/machine mises en œuvre. Le but visé est de dégager des recommandations pour la conception de logiciels permettant ces gestes d'appropriation du document numérique par le lecteur-interprète.

3.4 Domaine des pratiques scientifiques

Rapport PSI Pédagogue

L'équipe constituée par W. Turner, M.-J. Pierrat et M. de Saint Leger (LIMSI, IRESCO, DIST-CNRS) travaille dans le domaine des pratiques scientifiques et plus particulièrement dans le cadre de l'atelier sur l'importance des fonctionnalités annotatives associées à ces pratiques. Une enquête récemment réalisée par cette équipe montre que les sociologues français utilisent Internet pour la production collective de leurs connaissances disciplinaires. Ils travaillent de plus en plus dans des espaces scientifiques ouverts et sans murs : des « laboratoires de la recherche » (Turner, 1995). Internet leur sert non seulement pour accéder aux informations disponibles sur le Web mais également pour entrer directement en contact avec leurs collègues de la même spécialité ou de spécialités différentes. Comment décrire les pratiques informationnelles structurant l'activité collective d'un laboratoire ? Leur démarche expérimentale est organisée autour de trois questions qui sont au cœur d'un effort visant à soutenir informatiquement le développement des pratiques collectives distribuées sur Internet¹ et qui constituent autant de perspectives de recherches :

1. Les pratiques informationnelles ont une dimension éminemment cognitive, ce qui renvoie dans le cadre de ce travail à la construction des lexiques partagés. Comment les acteurs d'un laboratoire, qui travaillent dans des contextes géographiques, économiques et culturels hétérogènes, forgent-ils un vocabulaire commun ?
2. Les pratiques informationnelles ont également une dimension sociale, le problème étant celui de délimiter un espace d'interaction collective au sein d'un espace collaboratif virtuellement « infini ». Quelles sont les limites d'un réseau de recherche ? Qui doit participer à l'activité d'un laboratoire ? La station de travail lexical est structurée pour permettre l'étude de cette relation entre le social et le cognitif de par son usage des méta- données.
3. Enfin, les pratiques informationnelles ont une dimension organisationnelle : un travail lexical peut déboucher sur des propositions visant à tisser des liens de collaboration au sein d'un laboratoire.

Ce programme de travail est envisagé comme une contribution à une réflexion générale sur le rôle des professionnels de l'information dans le développement des recherches en laboratoire (Turner & Pierrat, 2003 ; Pierrat *et al.*, 2004).

¹ Pour avoir plus d'information sur le programme « Pratiques Collectives Distribuées » du LIMSI-CNRS : <<http://www.limsi.fr/Individu/turner/>>.

4 CONCLUSION

La diversité des recherches abordées au sein de l'atelier montre bien l'importance de la problématique de la gestion des annotations dans les documents pour l'action. Un des enseignements de l'atelier est la constatation selon laquelle la problématique de l'annotation ne concerne pas uniquement les documents professionnels au cœur des processus « métiers » majeurs des organisations comme le dossier patient ou les documents de conception, mais également des documents que l'on aurait pu penser être plus stabilisés, comme les articles de recherche ou les documents contenu dans des bibliothèques. La numérisation vient encore stimuler la seconde vie de ces documents en permettant à des chercheurs-lecteurs tout à la fois de prolonger le projet initial des auteurs en les articulant à leurs propres travaux et de les re-documentariser par le biais d'annotations spécifiquement documentaires visant à accroître le réseau de leurs bénéficiaires potentiels.

Dans le premier cas (annotation-fragment ou annotation-contributive), il s'agit d'associer au document source, sans bien sûr l'altérer, des fragments contributifs qui viennent éclairer ou orienter la lecture dans le contexte d'un nouveau projet d'écriture qui tout à la fois présente un point de vue synthétique sur une question et rassemble une collection de sources inédites à l'appui des thèses défendues. La documentarisation du fragment, qui le constitue en annotation, vient expliciter ses liens avec le corpus source, avec ses auteurs et avec leur problématique. Dans le second cas (annotation-index ou annotation-classificatoire), il s'agit de la possibilité donnée à des chercheurs travaillant sur des corpus de les enrichir « à la marge » pour les intégrer dans des collections constituées par leurs propres listes de référence elles-mêmes structurées selon un point de vue canonique associé à leur champ d'investigation (par exemple, une liste de liens vers des documents pertinents sur un sujet organisé en rubriques, en imaginant que la référence à la liste puisse elle-même devenir un attribut du document source partagé).

Sans se substituer au travail des professionnels de la documentation qui doivent rester les garants d'un standard de référence, cette re-documentarisation vient considérablement enrichir les possibilités d'accès aux corpus sources. De ce point de vue, les professionnels de la documentation trouveront dans ces systématiques, correspondant par exemple à des ontologies sémiotiques (Zacklad, 2005b), des outils permettant d'offrir de nouveaux services à leurs clients. Bien sûr, la condition de ces nouveaux usages des corpus sources dépend au préalable d'un accès le plus libre possible à ces derniers, une position défendue par

Rapport PSI Pédagogue

les défenseurs des archives ouvertes et systématisée sur le plan juridique par L. Lessing (Lessing, 2004).

Dans le contexte des DopA métiers, l'atelier souligne l'importance de nouveaux documents fragmentaires souvent constitués en dossier qui jouent un rôle majeur dans l'articulation des activités coopératives distribuées. Du point de vue des pratiques, les frontières entre les traitements de textes offrant des fonctionnalités annotatives, les forums de discussion, les wikis et les blogs tendent de plus en plus à s'atténuer, mais les logiciels gestionnaires de ces technologies ne sont encore que très faiblement inter-opérables. Les évolutions des pratiques plaident pour des transformations profondes des éditeurs de texte et la promotion de standards d'interopérabilité qui prennent sérieusement en compte les exigences de la coopération médiatisée. Par ailleurs, l'archivage de ces DopA et leur indexation selon des points de vue multiples, non seulement dans leur forme finale mais également à certains stades de leur évolution est un enjeu majeur de la problématique de la gestion des connaissances.

5 RÉFÉRENCES

- Barcellini F., Détienne F., Burkhardt J. M. & Sack W., A study of online discussions in an Open-Source community: Reconstructing thematic coherence and argumentation from quotation practices. *Communities and Technologies Conference, Milano (It.), June 2005*.
- Bénel A., *Consultation assistée par ordinateur de la documentation en Sciences Humaines : Considérations épistémologiques, solutions opératoires et applications à l'archéologie*, Thèse de doctorat en informatique, INSA de Lyon, décembre 2003.
- Blanco E. & Gardoni M., Supporting non structured graphical information in integrated design team. *Proceedings of the 13th International Conference on Engineering Design, ICED 01. Glasgow, UK, 21-23 August 2001*, p.11-19.
- Boujut J. F., User-defined annotations: artefacts for co-ordination and shared understanding in design teams. *Journal of Engineering Design, 14 (4), 2003*, p.409-419.
- Briet S., *Qu'est-ce que la documentation ?*, Paris : EDIT, 1951.
- Bringay S., Barry C. & Charlet J., Les documents et les annotations du dossier patient hospitalier. Numéro spécial de la revue *I3 Information, Interaction, Intelligence, Vol. 4, n°1, 2004*. p.191-211.
- Bringay S., Barry C. & Charlet J., Les annotations pour gérer les connaissances du dossier patient, *Actes de IC2005 Journées Françaises d'Ingénierie des Connaissances, Nice, 30 Mai – 3 Juin 2005a*. p.73-84.

Annotation et coopération

- Bringay S., Barry C. & Charlet J., Annotations for managing knowledge in the Electronic Health Record, A paraître dans les actes du Workshop *Knowledge management and organisational memories*, associé à l'IJCAI, Edinburgh (Scotland), 30 Juillet - 5 Aout 2005b.
- Brown J. S. & Duguid P., The Social Life of Documents. In E. Dyson (Ed.), *The Social Life of Documents*. New York : EDventure Holdings Inc., 1995. p.1-18.
- Buckland M. K., What is a "document"?, *Journal of the American Society for Information Science*, 48(9), 1997. p.804-809.
- Buckland, M. K., What is a "digital document"? *Document Numérique*, 2(2), . 1998, p.221-230.
- Carruthers M., *The Book of Memory: A Study of Memory in Medieval Culture*. New York: Cambridge University Press, 1990.
- Egyed-Zsigmond E., *Gestion des connaissances dans une base de documents multimédias*, Thèse de doctorat en informatique, INSA de Lyon, octobre 2003.
- Gasser L. & Ripoché G. Distributed Collective Practices and F/OSS Problem Management: Perspective and Methods, *Conférence "Coopération, Innovation & Technologie" (CITE2003)*, Troyes, 2003.
- Guibert S., Darses F. & Boujut J. F., Rôle des annotations dans l'élaboration collaborative d'une solution en conception mécanique, *Troisièmes Journées d'Etude en Psychologie Ergonomique EPIQUE'2005, Toulouse, 26-28 Septembre 2005*.
- Lessing L., *Free Culture*, The Penguin Press, 2004.
- Lortal G., Lewkowicz M. & Todiracu-Courtier A., Modélisation de l'activité d'annotation discursive pour la conception d'un collecticiel support à l'herméneutique, In : *Actes des 16ème journées francophones d'ingénierie des connaissances [IC 2005]*, Grenoble : PUG, 2005. p. 169-180.
- Maurer H. (Ed.), *HyperWave: The Next Generation Web Solution*, Addison-Wesley, 1996, 635 p.
- Mazhoud O., Pascual E. & Virbel J., Représentation et gestion d'annotations, In : *Hypertextes et hypermédias : Réalisations, outils et méthodes*, Techniques de l'Information, Paris : Hermès, 1995. p. 127-138.
- Pédauque R. T., *Document : forme, signe et médium, les re-formulations du numérique*, Working paper, 2003.
- Perlerin V., Sémantique légère pour le document : Assistance personnalisée pour l'accès au document et l'exploration de son contenu, In : F. Rastier (Ed.), *Texto ! décembre 2004*.
- Personal Brain: User Guide*, TheBrain Technologies Corporation, 2001, 99 p.
- Pierrat M.-J., de Saint Leger M. & Turner W., From information to knowledge management: a new challenge for information professionals, *19th International CoData Conference, Berlin, November 2004*.

Rapport PSI Pédagogue

- Prié Y., *Représentation de documents audiovisuels en Strates-Interconnectées par les Annotations pour l'exploitation contextuelle*, Thèse de doctorat en informatique, INSA de Lyon, décembre 1999.
- Tanguy L., Traitement automatique de la langue naturelle et Interprétation : Contribution à l'élaboration d'un modèle informatique de la Sémantique Interprétative, In : F. Rastier (Ed.), *Texto ! mars 1998*.
- Turner W. A., Les professionnels de l'information auront-ils une place dans les laboratoires de la recherche ? In : J.-M. Noyer (Ed.), « *Bibliométrie Scientométrie Infométrie* », *Revue Solaris*, 1995.
- Turner W. A. & Pierrat M.-J., Introduction au thème : « Collaboration chercheurs- documentalistes : De la recherche à l'analyse de l'information », *Les rencontres des professionnels de l'IST, Paris, 17-19, juin 2003*.
- Veron M., *Modélisation de la composante annotative dans les documents électroniques*, Mémoire de DEA RCFR INPT, Toulouse, Septembre 1998.
- Weick K. E., *The Social Psychology of Organizing*, New York : Random House, 1979.
- Zacklad M., Lewkowicz M., Boujut J. F., Darses F. & Détienne F., Formes et gestion des annotations numériques collectives en ingénierie collaborative. In R. Dieng (Ed.), *Actes des 14èmes journées francophones d'Ingénierie des Connaissances [IC'03]*, Grenoble : PUG, 2003, p.207-224.
- Zacklad M., Communities of Action: A Cognitive and Social Approach to the Design of CSCW Systems. In : *Proceedings of GROUP'2003, Sanibel Island (Florida)*, 2003. p.190-197.
- Zacklad M., Processus de documentarisation dans les Documents pour l'Action (DopA) : statut des annotations et technologies de la coopération associées, In : *Le numérique : Impact sur le cycle de vie du document pour une analyse interdisciplinaire, Montréal, 13-15 Octobre 2004*. 19 p.
- Zacklad M., Innovation et création de valeur dans les communautés d'action, in P. Lorino et R. Teulier (Ed.) *Connaissance, Activité, Organisation*, Maspéro, 2005a.
- Zacklad M., Vers le Web Socio Sémantique, In : *Actes des 16ème journées francophones d'ingénierie des connaissances [IC 2005]*, Grenoble : PUG, 2005b.