

Processus de documentarisation dans les Documents pour l'Action (DopA) : statut des annotations et technologies de la coopération associées

Manuel Zacklad

ISTIT/Tech-CICO (FRE CNRS 2732)
Technologie de la Coopération pour l'Innovation et le
Changement Organisationnel

Université de Technologie de Troyes

Problématique générale

- Tech-CICO (Technologie de la Coopération pour l'Innovation et le Changement Organisationnel):
équipe de recherche pluridisciplinaire sur la gestion des connaissances et les activités coopératives
- Relation avec les réflexions documentaires du RTP-DOC mais en relation avec des documents plus « triviaux » : KM plus dans l'histoire immédiate que la grande histoire

Introduction: usages collectifs du Web et documents

- L'essor du Web entraîne 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 non (activité d'associations militantes ou de communauté d'intérêt partageant une passion commune).
- Nouveaux usages : 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 Blog.

DopA et cycle de vie

- Notion de Document pour l'Action (DopA) qui met l'accent sur les pratiques d'hyperédaction et l'hyperfragmentation associée.
- Pratiques qui questionnent à plusieurs titres les nouvelles « théories du document » et notamment la problématique de la gestion de son **cycle de vie**
- Son existence publique peut commencer avant qu'il soit achevé et il est en mesure de rassembler de nouvelles contributions sous la forme d'annotations

Plan

- **Production sémiotique** dans le cadre d'une transaction communicationnelle
- Pallier à la distribution spatio-socio-temporelle de ces transactions la **documentarisation**
- Analyse de la documentarisation à travers plusieurs notions: distinctions transcription vs enregistrement, double articulation documentaire externe vs interne du document, les liaisons mécanique et organique entre les fragments du document.
- Définition des **Document pour l'Action** (DopA) et leurs caractéristiques.
- Coopérer dans les DopA grâce aux **annotations**
- Analyse de technologies de la coopération et en particulier les **forums de discussion** et rôle de la codification

Productions sémiotiques dans les transactions communicationnelles

Le document objet de transactions communicationnelles

- Dans les termes de la *théorie des transactions communicationnelles symboliques*, le document s'inscrit dans un procès de communication qui implique en principe une forme d'échange (c'est la raison pour laquelle les romanciers reçoivent un courrier important de la part de leurs « lecteurs »)
- Dans une transaction communicationnelle il y a échange de reconnaissance mutuelle (constitution des **selfs**) et de connaissances sur le monde (production d'une **œuvre**)

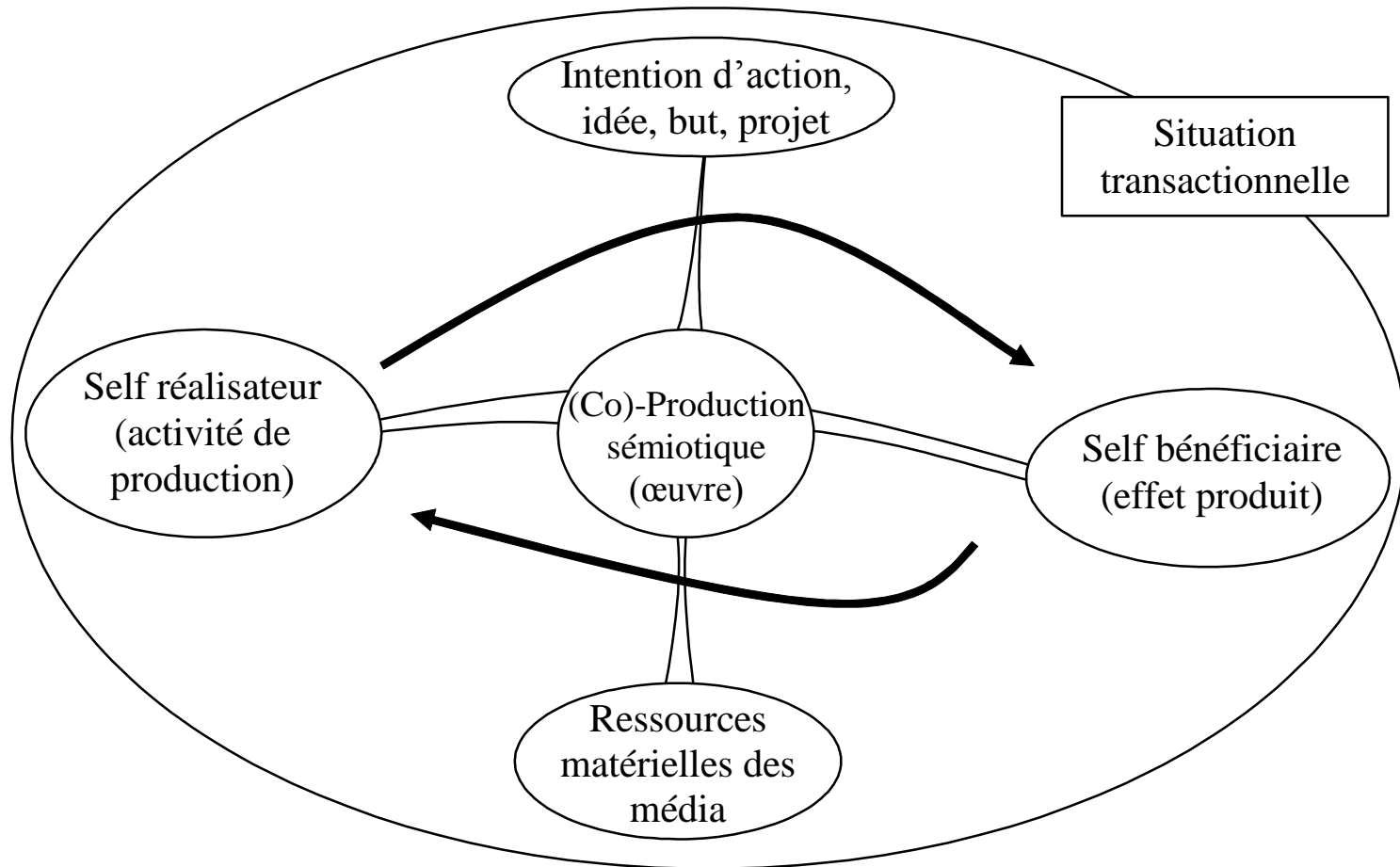
Importance des médias dans les transactions

- Dans la transaction communicationnelle l'acteur n'agit pas *directement* sur l'interlocuteur
- Il le fait via une « production sémiotique » impliquant une activité de mise en forme qui nécessite une activité de réception de la part du bénéficiaire de la transaction
- La production sémiotique peut être de nature très diverse : geste, énoncé oral, écrit, film...

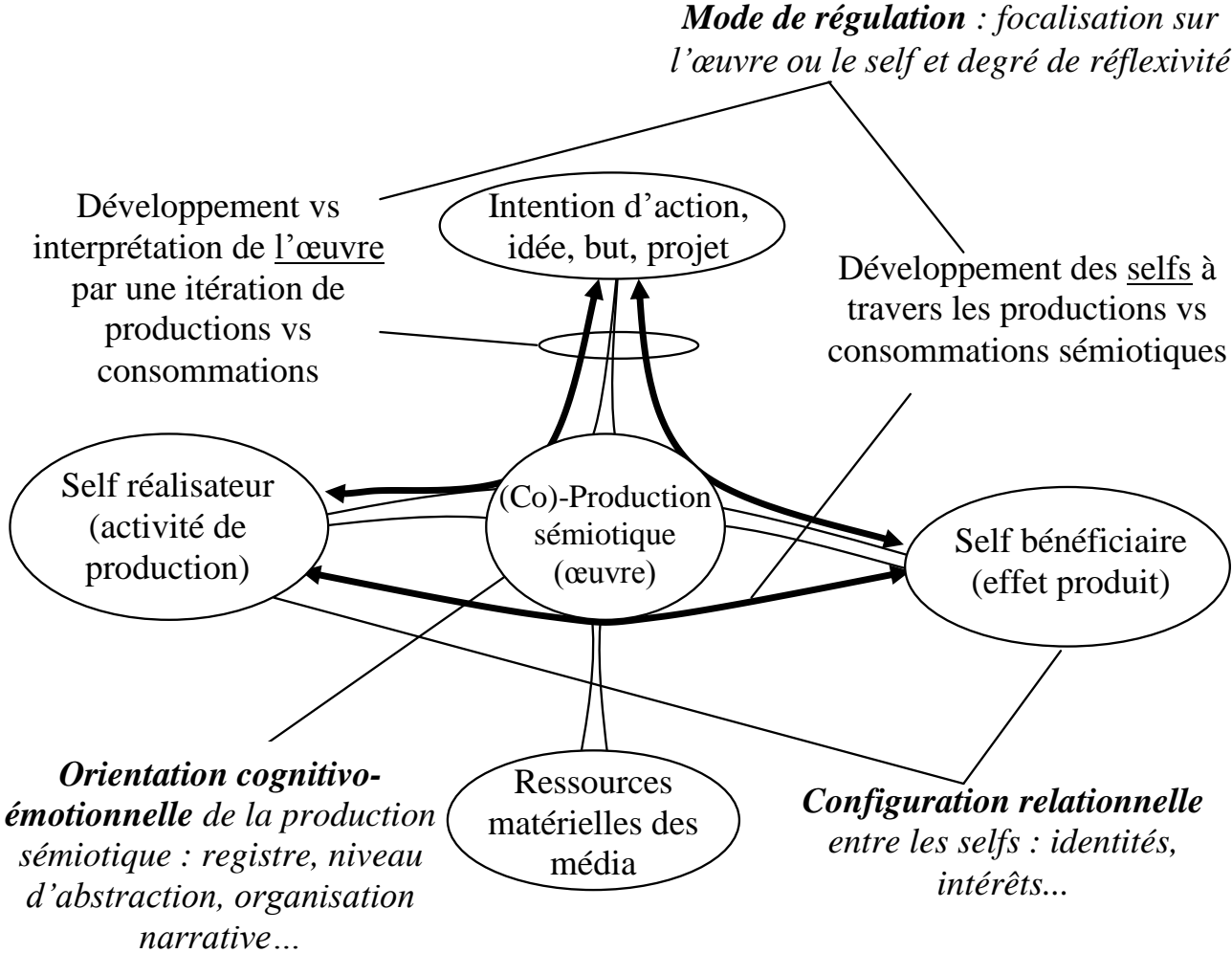
Production sémiotique élaborée dans une situation transactionnelle

- Les composantes de la *situation transactionnelle* (qui peut elle-même se décomposer dans certains cas particuliers en situation de production sémiotique et situation de réception) comprennent :
 - Des **selfs** en présence qui sont les acteurs de la transaction et qui constituent :
 - un ou plusieurs **réalisateurs** (individuels ou collectifs) ;
 - un ou plusieurs **bénéficiaires** (individuels ou collectifs) ;
 - une *production sémiotique* ;
 - des *paramètres de la situation transactionnelle*

Schématisation des composantes de la transaction communicationnelle



Dimensions pour la caractérisation des transactions communicationnelles symboliques



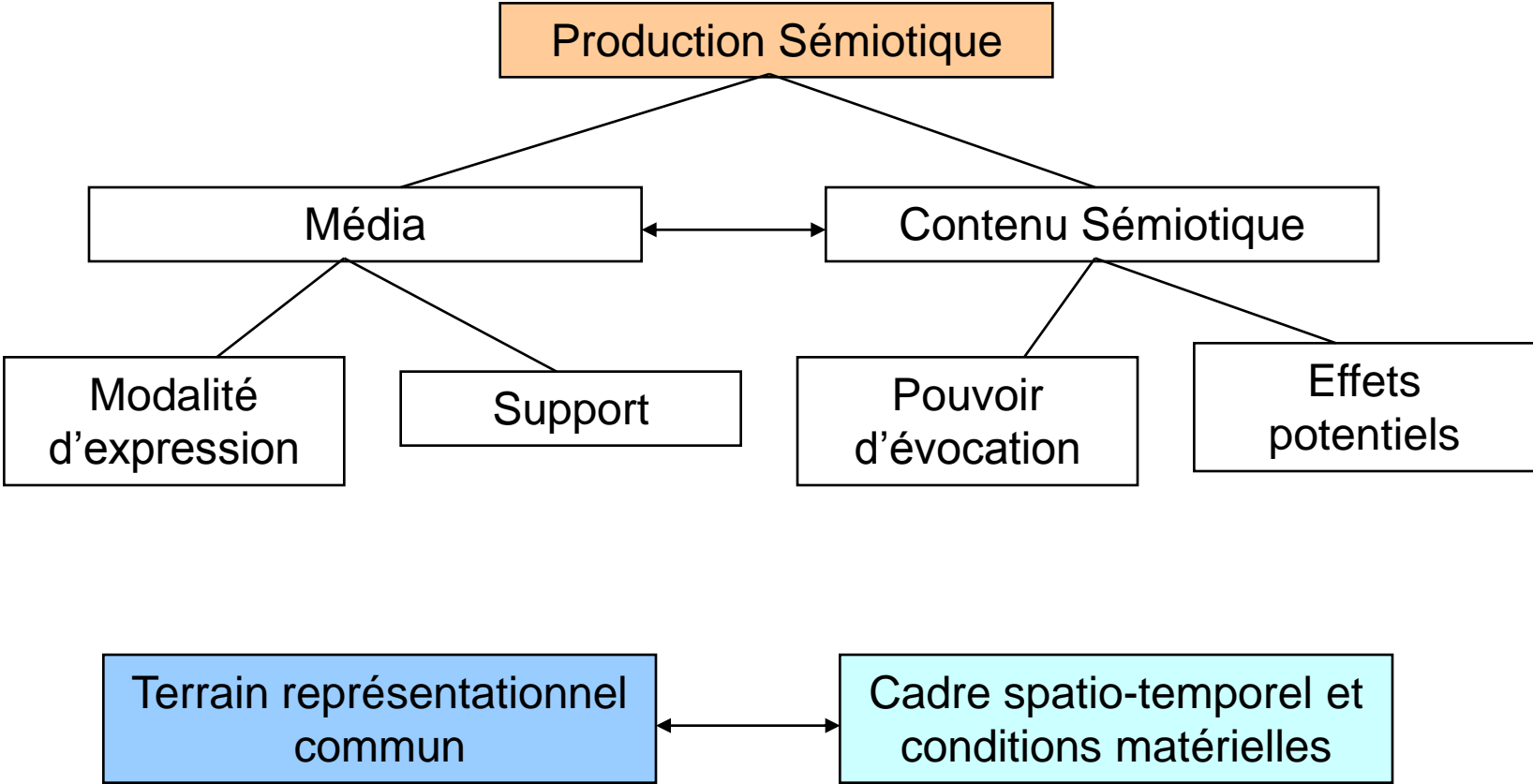
Production sémiotique : média et contenu sémiotique

- **Média** qui regroupe lui-même :
 - le choix d'une **modalité d'expression** (langagière orale, scripturale, gestuelle, filmique, etc...), au sein de laquelle est conventionnellement déterminé un système de signes permettant des agencements syntagmatiques et paradigmatiques ;
 - le choix d'un **support matériel** associé à cette modalité d'expression (type de vocalisation, amplitude des gestes, nature du support pour un média scriptural, etc...) adapté aux caractéristiques de la modalité d'expression ;

Contenu sémiotique de la transaction « véhiculé » par le média

- **Contenu sémiotique** appréhendé selon deux facettes :
 - son pouvoir d'évocation qui est sa capacité à susciter des *représentations partagées* (le *pouvoir d'évocation* correspond à la facette traditionnellement étudiée par la sémantique) ;
 - ses effets potentiels psychiques et sociaux qui correspondent aux *conséquences possibles* de l'évocation de certaines représentations qui permettent d'attester de l'effectivité de la communication (les *effets potentiels* correspondent à la facette traditionnellement étudiée par la pragmatique linguistique).

SITUATION TRANSACTIONNELLE



Distribuer les transactions
communicationnelles par la
Documentarisation

Stratégies pour la distribution des transactions communicationnelles

- Passage de l'interaction de synchrone et de face à face dans des communautés restreintes à une distribution élargie
- Quatre stratégies :
 - Normalisation de la situation transactionnelle
 - Ritualisation mnémotechnique
 - Formalisation de l'expression
 - Documentarisation
- Stratégies non indépendantes les unes des autres et qui se renforcent

Passage en revue rapide des différentes stratégies de distribution

- Normalisation de la situation transactionnelle
 - Standardisation du projet commun, des relations sociales, du cadre matériel, etc... : *contractualisation, homogénéisation de l'architecture dans des agences ouvertes au public*
- Ritualisation mnémotechnique
 - Récitation, techniques d'association mentale, utilisation d'artefact externes stimulant l'évocation des représentations mentales... : *formation, campagnes de communication...*
- Formalisation de l'expression
 - Travail sur cette composante du média (dimension syntagmatique et paradigmatic) : *langage non ambigu, terminologie...*
- Documentarisation
 - Travail sur la composante « **support** » du média et investissement documentaire visant à le doter d'attributs permettant sa ré-exploitation : *utilisation des supports écrits, TIC*

Documentarisation

- Deux types de supports matériels : éphémère ou pérenne.
 - Les *supports éphémères* sont caractérisés par le fait que les stimuli qu'ils véhiculent ne peuvent exercer qu'un effet transitoire sur les organes des sens des récepteurs notamment du fait que ceux-ci ne peuvent exercer de contrôle direct sur le déclenchement et la reproduction des stimuli
 - Par contraste, les *supports pérennes* sont susceptibles d'exercer un effet réitéré sur les organes des sens des récepteurs notamment grâce au fait que ceux-ci peuvent exercer un contrôle sur le déclenchement des stimuli

Définition de la stratégie de documentarisation

- Consiste à *transcrire ou à enregistrer les contenus sémiotiques sur des supports pérennes puis à doter ces supports d'attributs spécifiques permettant :*
- *de faciliter leur gestion « externe » parmi d'autres supports;*
- *de faciliter leur gestion « interne » :*
 - *manipulation physique, condition d'une navigation sémantique à l'intérieur du contenu sémiotique*
 - *orientation des récepteurs, mais également de plus en plus des réalisateurs eux-mêmes, à l'intérieur du support en définissant une ou plusieurs cartographies des contenus sémiotiques susceptibles de guider la navigation sémantique.*

Transformation du support

- Il existe deux méthodes permettant de véhiculer une production sémiotique sur un support pérenne :
 - **Transcription** : le support de signes graphiques dans des logiques de figuration ou d'équivalence phonétique, par exemple. Elle implique une transposition du contenu sémiotique dans une modalité d'expression nouvelle.
 - **L'enregistrement** : via des dispositifs le plus souvent aujourd'hui électroniques

Investissement documentaire de la transcription ou de l'enregistrement

- Transcription et enregistrement sont des conditions nécessaires mais non suffisantes de la documentarisation
- Une prise de notes ou un enregistrement ponctuel visant à faciliter une activité sémiotique dans une situation transactionnelle donnée pourront avoir une utilité locale certaine.
- 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

Analyse du document sous l'angle de la documentarisation

Définition du document

- Le document désigne une *production sémiotique transcrite ou enregistrée sur un support pérenne qui est équipée d'attributs spécifiques, plus ou moins codifiés, visant à faciliter les pratiques liées à son exploitation ultérieure dans le cadre de la préservation de transactions communicationnelles distribuées en articulant sémantiquement le contenu vis-à-vis d'autres ressources externes et dans son organisation interne.*
- Ces attributs visent à 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.

Articulation sémantique externe et interne du document

- Retrouver le document par parmi d'autres documents parmi lesquels il est stocké à l'intérieur de *bibliothèques, d'archives, d'armoire de classement, de base de données administratives, techniques ou documentaires* de nature diverses : **attributs permettant l'articulation sémantique externe** du document
- Retrouver une sous-production sémiotique donnée « à l'intérieur » du document : **articulation sémantique interne** de type organique ou mécanique (document composé de transactions locales déployant un univers d'actant et d'objets)

Lier des fragments documentaires à l'intérieur d'un support de taille importante : **la liaison des fragments**

- L'utilisation d'un support pérenne pour médiatiser la circulation des contenus sémiotiques va également permettre de transcrire ou d'enregistrer sur le même support un nombre très important de signes et donc de transactions ou de sous transactions partielles: plurivocité, pluritextualité
- Une sous-production sémiotique transcrite ou enregistrée sur un support constitue un fragment
- On distingue deux types de liaison sémantique : **organique** (plus naturelle et tacite) et **mécanique** (plus artificielle et formelle)

Liaison organique

- Les liaisons organiques sont celles qui correspondent à l'enchaînement naturel des fragments qui exploite les différentes logiques de planification associées au projet transactionnel du réalisateur : temporelle, structurelle, fonctionnelle, etc.
- Elles exploitent également toutes les relations de présupposition autorisées par le partage d'un terrain représentationnel commun, modifié et élargi au fur et à mesure du déroulement de la transaction.
- Ce découpage est la fois celui qui peut sembler le plus « naturel » et celui qui semble fournir une logique d'accès implicite au contenu sémiotique.

Liaison mécanique

- Le système d'orientation mécanique vient doubler le précédent en fournissant un système d'orientation supplémentaire qui soit, explicite le projet navigationnel initial (sous-titrage et table des matières), soit vient proposer des projets navigationnels a priori plus inattendus (index, tables de référence...).
- Dans ce dernier cas, il s'appuie le plus souvent sur un ensemble de ressources terminologiques et ontologiques organisées selon une logique différente de celle de la « table des matières », à la fois plus générale et plus systématique ou exploitant un point de vue très particulier, comme le nom des auteurs cités.
- Sa constitution demande toujours un investissement significatif mais peut conférer au document une valeur ajoutée importante correspondant à une plus grande polyvalence transactionnelle.

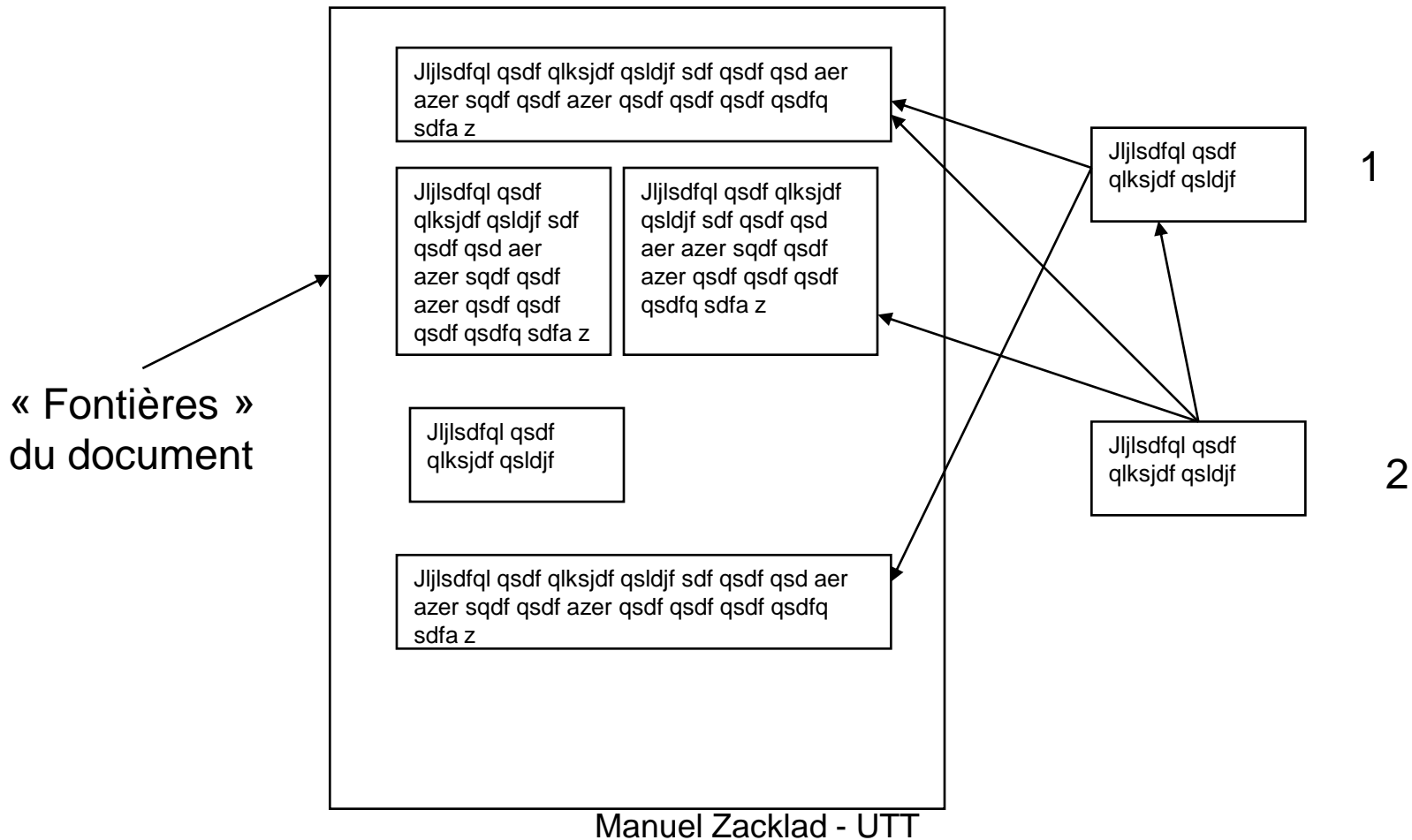
Document lisse

- Quand un document apparaît comme « lisse », semblant écrit d'un seul tenant, c'est que les différentes prises en charges énonciatives correspondant à des sous-transactions communicationnelles sont implicites.
- Les relations sémantiques entre celles-ci correspondant à une articulation organique dans laquelle les présupposés associés aux terrains représentationnels communs permettent un enchaînement naturel des « arguments ».
- Chaque technologie littéraire correspondant à un « genre » privilégie une modalité d'articulation organique spécifique de manière implicite.

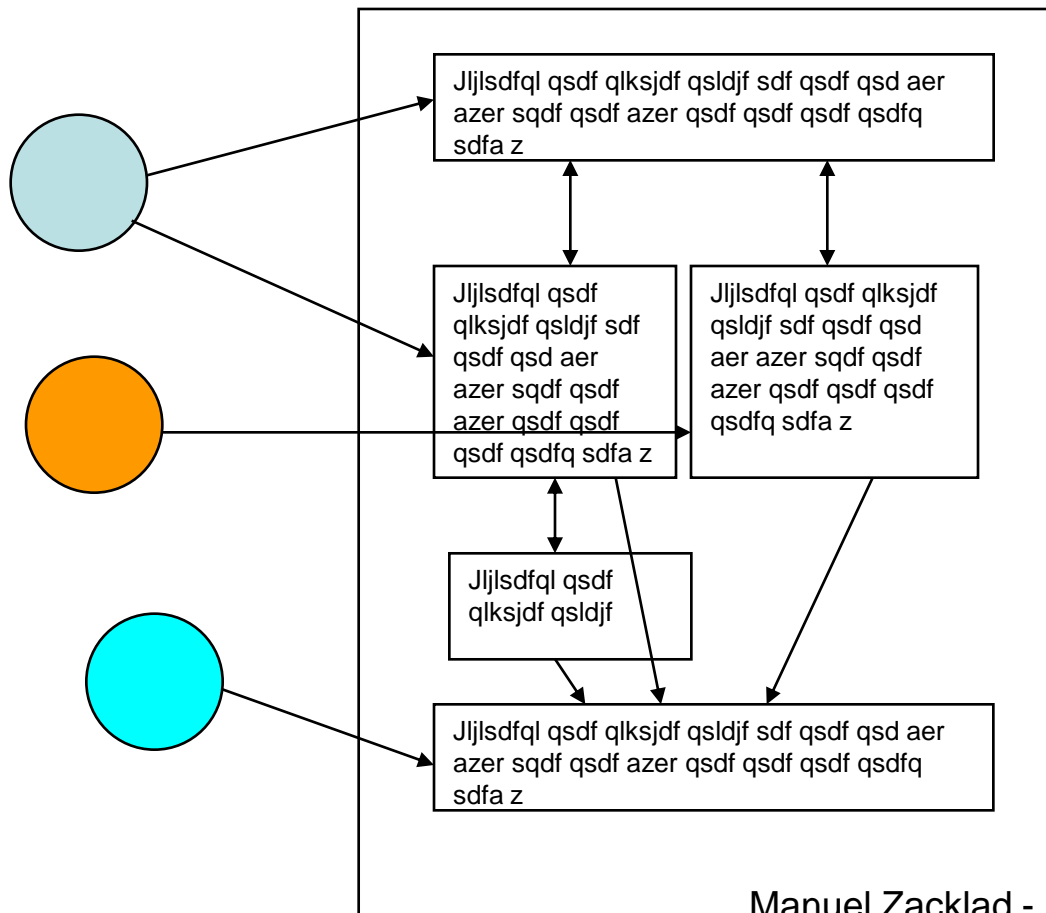
Document fragmenté

- Inversement d'autres documents apparaissent comme beaucoup plus « fragmentés ».
- Les sous-production sémiotiques sont mécaniquement articulées aux précédentes par le biais d'un système de numérotation, d'un sous-titrage, d'un ensemble d'attributs spécifiques visant à expliciter le statut des fragments dans le document et leur relation avec les fragments environnants.
- Les liaisons mécanique pourront elles-mêmes apparaître comme « faibles » eu égard à la cohérence discursive.
- Cette fragmentation pourra être extrême dans une classe de documents jouant un rôle essentiel dans la coordination des transactions communicationnelles distribuées les « documents pour l'action ».

Annotation et argumentation avec support graphique

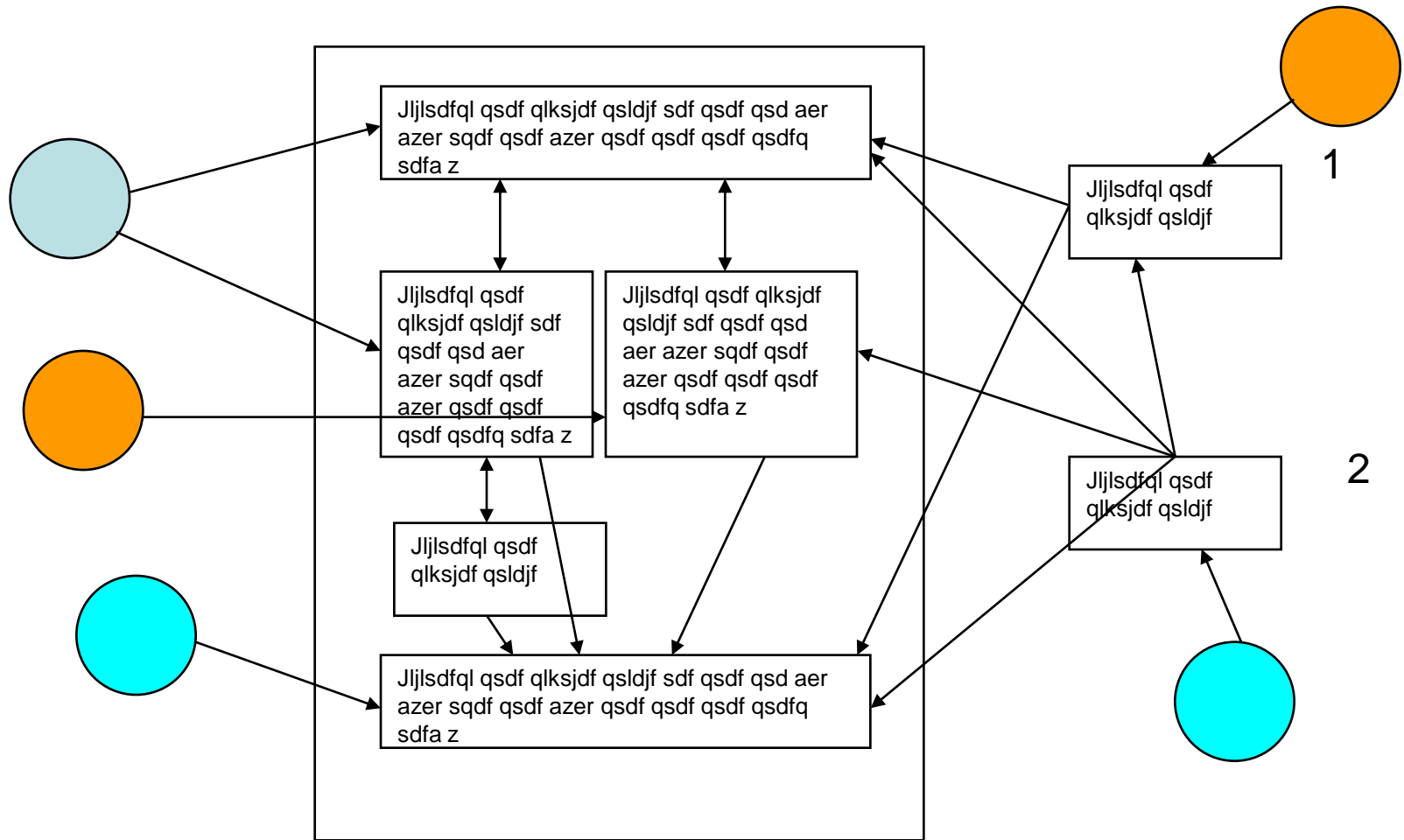


Formulaire : mécanisme de coordination (K. Schmidt)

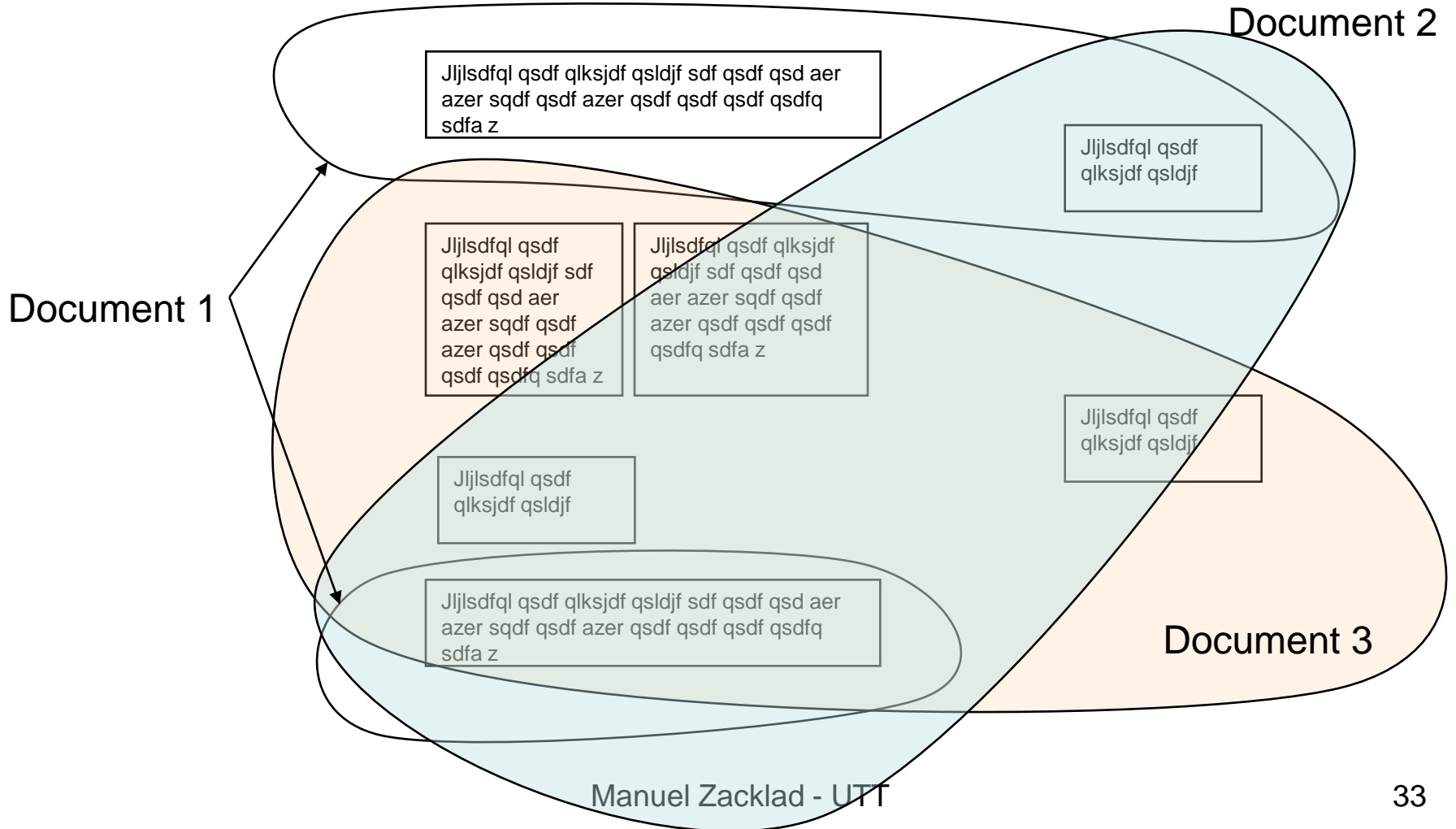


« Un mécanisme de coordination est un construit consistant en un protocole de coordination (un ensemble de procédures et de conventions définissant les modalités d'articulation d'activités interdépendantes complexes) et en un artefact (un construit symbolique permanent) dans lequel le protocole est matérialisé. » (Schmidt 1996)

Formulaire + annotation : mécanisme de coordination permettant la gestion d'interactions émergentes



Les frontières « naturelles » du document s'effacent... celui-ci devient malléable



Présentation d'une classe de documents fragmentés: les Documents pour l'Action (DopA)

Caractéristiques des Documents pour l'Action (DopA)

- Désigne une classe de documents, numériques rassemblant, au fur et à mesure de leur élaboration, les contributions d'un collectif engagé dans une activité commune
- Notion de « document pour l'action » qui peut être prolongée par celle de « dossier pour l'action », hybride complexe à la fois:
 1. soutien à l'activité collective pendant sa rédaction;
 2. preuve et témoin des décisions ou des analyses dont il acte l'existence;
 3. mémoire des transactions associées aux processus de résolution de problème et de délibération dont il est le résultat;
 4. offrant des informations exploitables par les processus visant à l'indexer, le rechercher et l'exploiter.

Définition des DopA

- *Documents qui soutiennent de façon évolutive les transactions communicationnelles associées aux actions collectives*
 1. **Inachèvement prolongé** : ils possèdent un statut d'inachèvement prolongé pendant la phase de réalisation du document ou du dossier phase durant laquelle nous les appellerons des *DopA en évolution* (vs *DopA stabilisés*) ;
 2. **Pérennité** : ils possèdent une certaine pérennité associée aux engagements des acteurs à l'égard des contenus véhiculés d'où des enjeux liés à leur stockage et leur indexation ;
 3. **Fragmentation** : au moins pendant leur phase évolutive ils sont composés de plusieurs fragments qui ne peuvent être fusionnés ;
 4. **Rapport complexe aux contributeurs** : les différents fragments sémiotiques sont le plus souvent produit par des contributeurs multiples (plurivocité, pluri textualité...) qui peuvent avoir des statuts variables vis-à-vis du contenu (droits, engagement, responsabilités...) ;
 5. **Rapport non trivial à la situation transactionnelle et aux autres fragments** : modalisation, incertitude, relation argumentative vis-à-vis des autres fragments...

Une diversité de contextes applicatifs des DopA

- Ils couvrent un très grand nombre de contextes professionnels, parmi lesquels on peut citer :
 - Les documents de conception en ingénierie (mécanique, logiciel...)
 - Les dossiers patients en médecine ;
 - Le document contractuel dans un contexte d'affaires, qui passe du stade de proposition commerciale au stade de contrat en bonne et due forme ;
 - Un dossier qualité ;
 - Un rapport d'étude dans le conseil en management ;
 - Un forum d'échange dans le domaine du logiciel libre.

Une variété de technologies « littéraires » et documentaires

- Degré variable de standardisation de leur structure qui reflète la plus ou moins grande standardisation de la situation transactionnelle :
 - Formulaire
 - Plan type (p.e. modèle systémique en ingénierie, ou contrat juridique)
 - Convention d'écriture (style) p.e. descriptive ou argumentative normée
 - Système de question/réponse de type forum
 - Document narratif (visant à masquer le plan)
- Par ailleurs au sein de ces documents l'expression peut être plus ou moins formelle

Coopérer à travers un DopA
via des Annotations

Pallier via le DopA à la flexibilité des transactions communicationnelles dans les interactions polylogales associées aux activités coopératives

- La difficulté à laquelle sont confrontés les acteurs qui coopèrent via le DopA (par exemple un forum) est le manque d'information sur les paramètres du contexte transactionnel associé à l'offre de transaction par comparaison avec les polylogues synchrones
- Si les fragments déposés sur le support ne sont pas suffisamment liés au reste du document au moment de leur dépôt la reprise du DopA est impossible (puisque les liaisons organiques qui supposeraient un modèle commun font défaut).
- Il est donc nécessaire *d'explicitier sur le support du fragment certains paramètres de la situation transactionnelle* pris en compte par les participants c.à.d documentariser les fragments sémiotiques

Annotation & activité annotative

- *L'activité annotative résulte d'un effort de documentarisation d'un fragment visant à le lier mécaniquement avec d'autres composantes du contenu sémiotique et l'annotation est un fragment documentarisé lié au corps du document ou à d'autres annotations*
- Les transactions communicationnelles annotatives sont la transposition dans un contexte documentaire des transactions communicationnelles polylogales en situation de face à face

Liaison sémantique codifiée des annotations

- Bien que les forums soit présentés comme des systèmes de question-réponse à distance basés sur un point de vue dialogique, des études détaillées montrent qu'ils ne correspondent pas vraiment aux « modèles communicationnels » traditionnels des polylogues en présentiel.
- *Les forums sont en fait des technologies documentaires annotatives qui soutiennent les transactions distribuées en proposant **une liaison semi-automatique des fragments par le biais d'attributs codifiés** affectés à chaque contribution pour faciliter sa documentarisation.*
- En fait la possibilité donnée d'intervenir à tout moment sur un contenu sémiotique même « ancien » confirme qu'ils correspondent d'avantage à des transactions annotatives.

Limites de la codification semi-automatique actuelle et projet socio-informatique

- Malheureusement les attributs codifiés fournis par le dispositif sont loin d'être suffisant, ce qui rend la documentarisation du forum incomplète pour des usages distribués selon une temporalité importante (et une socialité hétérogène).
- Les activités interprétatives des relecteurs restent essentielles pour accéder au contenu sémiotique

Un projet de recherche socio-informatique & Ingénierie des connaissances

- Poursuite des recherches relevant du CSCW et de l'ingénierie des connaissances pour la mise à disposition de systèmes de **ressources terminologiques et ontologiques** plus adaptés à la codification des transactions documentaires selon les différentes dimensions de l'activité collective : cognitive, relationnelle...
- Définition d'un standard relevant du **Web Socio Sémantique** (HyperTopic) permettant de représenter des ressources de manière évolutive et autant que possible proche des attentes des participants basées sur une forme de « négociation sectorielle » du sens basée sur les différents **points de vue** en présence.

Le Web Socio Sémantique (W2S) et HyperTopic

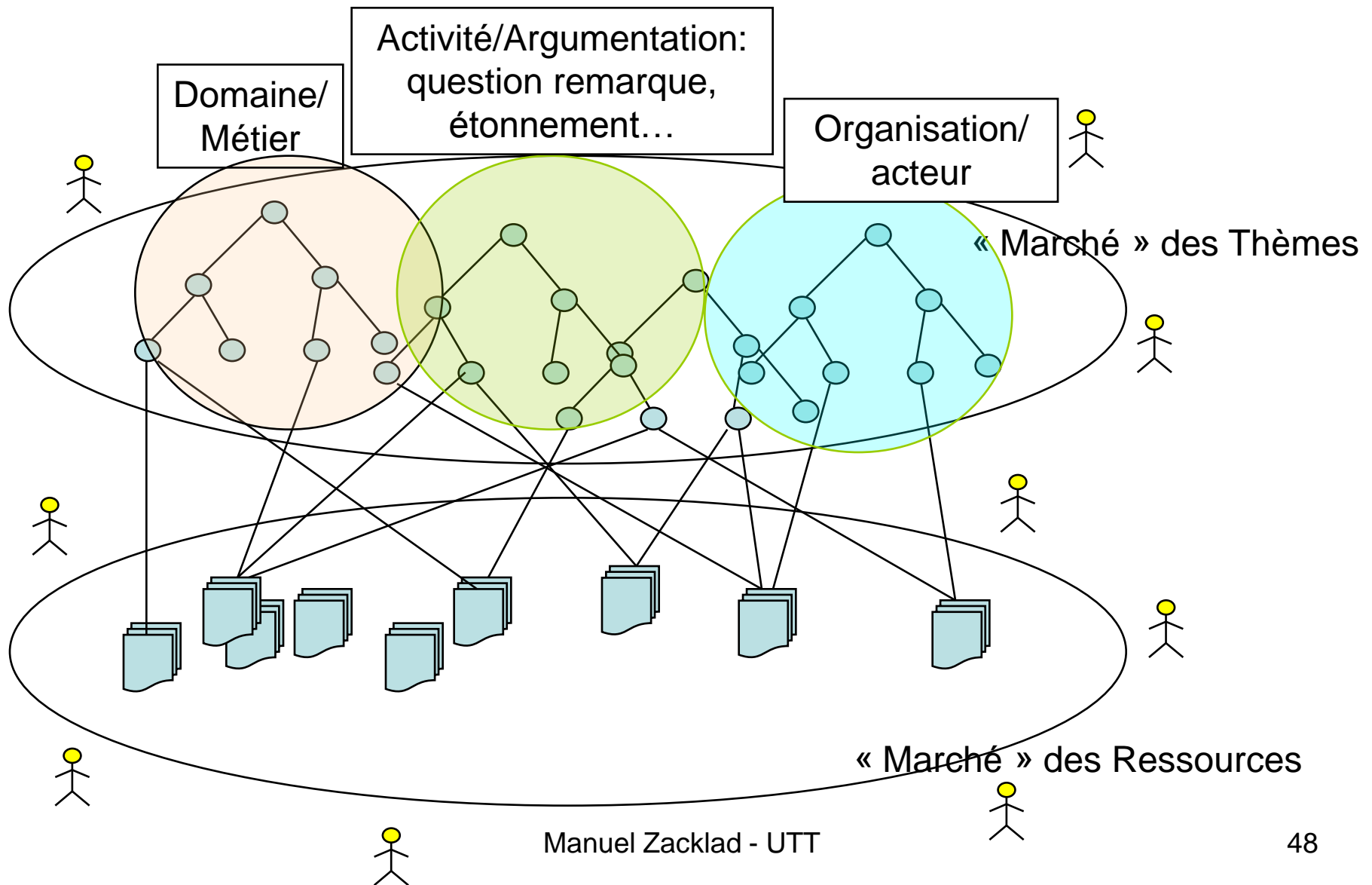
Un nouveau standard de modélisation pour le W2S

- Modéliser l'information (dans son sens « documentaire ») et son rôle dans une activité transactionnelle plutôt que « la connaissance en soi »
- Une modélisation qui s'appuie sur les pratiques sociolinguistiques de la dénomination au sein de communautés de parole et d'action et pas sur l'identification de « concepts » universels
- Un langage qui privilégie l'inférence heuristique sur le calcul formel déterministe (ou aux requêtes dans des bases de données bien structurées)

L'importance des « ontologies sémiotiques »

- Dans ce contexte les « ontologies formelles » (universelles, stables, entièrement calculables...), ne sont pas adaptées
- Recours à des « ontologies » sémiotiques permettant d'assurer des « traductions » entre domaines de spécialité et de refléter des consensus
- Importance de l'approche « multi-point de vue », où chaque point de vue représente une facette particulière

Des points de vue structurant et complémentaires...



Quelques liens...

- “Distributed Collective Practices: Building New Directions for Infrastructural Studies”. CSCW Workshop, Saturday, November 6th 2004 - organisateurs :
 - Geoffrey Bowker (University of California at San Diego)
 - Les Gasser (University of Illinois at Urbana Champaign),
 - Helena Karasti (University of Oulu),
 - Kjeld Schmidt (University of Copenhagen),
 - William Turner (LIMSI -CNRS),
 - Manuel Zacklad (Université de Technologie de Troyes)
- <http://www.acm.org/cscw2004/>
- www.yeposs.org - site de démonstration « Yellow pages for open source software » basé sur le langage HyperTopic développé à Tech-CICO et la plate-forme Agorae

Questions ?