



Disponible en ligne sur [www.sciencedirect.com](http://www.sciencedirect.com)



ScienceDirect

Sociologie du travail xxx (2008) xxx–xxx

SOCIOLOGIE  
DU TRAVAIL

<http://france.elsevier.com/direct/SOCTRA/>

# Les usages épistémiques des réseaux de communication électronique : Le cas de l'*Open-Source*

## The epistemic uses of electronic communication networks: Open source

Bernard Conein<sup>a,\*</sup>, Matthieu Latapy<sup>b</sup>

<sup>a</sup> UFR LASH, LASMIC, université de Nice Sophia-Antipolis, BP 209, 98, boulevard Édouard-Herriot, 06204 Nice cedex 3, France

<sup>b</sup> LIP6, CNRS, université Paris-6, 104, avenue du Président-Kennedy, 75016 Paris, France

### Résumé

L'objet de cet article est d'étendre la notion de cognition distribuée à l'analyse des réseaux de communication sur Internet pour obtenir une meilleure compréhension du rôle joué par la communication électronique dans l'acquisition et la construction de connaissance. Le cas retenu concerne une liste de diffusion destinée aux usagers du logiciel libre, *Debian.user.french*. Dans cet article, il est suggéré que les réseaux de communication, agencés en fils de discussion, gouvernent la distribution des connaissances au sein de la liste. La structure des fils de discussion exerce une contrainte sur la transmission des connaissances. Il est montré que les fils de discussion donnent naissance à deux types de réseaux épistémiques : des réseaux de conseil et des réseaux de discussion. Les réseaux électroniques de communication facilitent la cohabitation au sein d'une même communauté entre ces deux types de réseaux. L'existence d'un mécanisme de mise en discussion des conseils permet la sélection des informateurs réputés et l'émergence de structures collaboratives à partir des relations de conseil. Cette cohabitation d'usages épistémiques variés caractériserait les communautés *Open-Source* car elle conjugue élaboration collaborative de la connaissance et transmission instructionnelle aux apprentis.

© 2008 Publié par Elsevier Masson SAS.

### Abstract

The concept of "distributed cognition" is applied to analyzing communication networks on Internet in order to better understand the role of electronic communications in acquiring and producing knowledge. This case study has as its basis a mailing list to the users of the free software "Debian.user.french". The communication

\* Auteur correspondant.

Adresses e-mail : [conein@ehess.fr](mailto:conein@ehess.fr) (B. Conein), [matthieu.latapy@lip6.fr](mailto:matthieu.latapy@lip6.fr) (M. Latapy).

networks that emerged out of discussions on “threads” determined how knowledge was distributed among persons on the list. The structure of these threads limited the transmission of knowledge and gave rise to two types of epistemic networks: *counseling* networks and discussion networks. Electronic communication networks facilitate the coexistence of these two types in a single community. Procedures that lead to a discussion of *counseling* serve to select informants with a reputation and to develop cooperative structures out of the relationships formed while giving advice. This coexistence of epistemic uses characterizes open-source communities, since it combines the cooperative production of a body of knowledge and the transmission of instructions to learners.

© 2008 Publié par Elsevier Masson SAS.

**Mots clés :** Cognition distribuée ; Réseaux épistémiques ; Interaction ; Internet ; Fils de discussion ; Logiciel libre ; Communautés *Open-Source*

**Keywords:** Distributed cognition; Epistemic networks; Interactions; Internet; Discussion threads/forums; Free software; Open-Source communities

Une idée centrale proposée par l’approche de la cognition distribuée serait que la cognition individuelle<sup>1</sup> ne peut, et n’a pu, évoluer que grâce aux contributions coordonnées des artefacts informationnels et de la coopération sociale. Les groupes sociaux, comme les artefacts agissent alors comme des aides cognitives externes (Clark, 1997 ; Clark et Chalmers, 1998 ; Hutchins, 1995a ; Kirsh, 2006) ou des assistances épistémiques (Goldman, 1999). Mais cette contribution des artefacts ou des groupes sociaux à la cognition peut être comprise de plusieurs manières selon les propriétés que possèdent les artefacts eux-mêmes et selon les usages qui en sont faits. Dans cet article, nous nous intéresserons aux usages « épistémiques » des listes de diffusion par des utilisateurs du logiciel libre. Le cas retenu concerne les échanges de connaissances sur la liste *Debian.user.french* pendant une période de dix ans<sup>2</sup>.

L’objet de la recherche est de comprendre comment les réseaux de communication électronique favorisent des processus collaboratifs d’acquisition de connaissances. Nous suggérons que le développement des technologies numériques réticulaires (*Network technologies*) qui conjuguent aide cognitive et support à la communication favorise un usage épistémique des artefacts cognitifs où la diffusion et l’archivage de l’information sont orientés vers une production collaborative de la connaissance. En particulier, nous montrerons comment la structure des fils de discussion contribue à faire émerger des relations de discussion à partir des relations de conseil.

Le travail présenté ne porte pas sur le fonctionnement des communautés numériques du logiciel libre, ce qui a déjà été fait, ailleurs, dans d’autres contributions (Auray, 2005 ; Conein, 2004a ; Foray et Zimmerman, 2001 ; Lakhani et von Hippel, 2000). Le but est plus modeste : proposer une méthode d’analyse pour mettre en valeur la façon dont la distribution de la connaissance est affectée par la structure des réseaux de communication électronique. La compréhension du phénomène du logiciel libre nécessiterait de plus de soulever d’autres questions, comme par exemple

<sup>1</sup> Selon R. Giere et B. Moffatt, la cognition distribuée s’avère d’abord pertinente lorsqu’une tâche ne peut pas être accomplie par un individu seul : « les systèmes distribués rendent possible des acquisitions de connaissances impossibles à réaliser par un individu seul ou même par un groupe d’individus » (Giere et Moffatt, 2003).

<sup>2</sup> Deux études sur la communauté des debianistes ont été successivement menées, une première dans le cadre du projet CNRS « Société de l’Information, 2002 à 2004 » avec N. Auray (institut TELECOM)<sup>1</sup>, R. Dorat (LIFL), M. Latapy (LIP6) et une seconde dans le cadre de l’ANR « Autograph, 2005 à 2007 » avec N. Auray (institut TELECOM)<sup>2</sup> et M. Latapy (LIP6).

montrer comment l'émergence des communautés numériques est favorisée par une économie des biens informationnels (Gensollen, 2004). Cependant, cet article peut contribuer indirectement à la question de l'émergence des communautés, en soulignant le rôle des structures locales de communication en réseau.

## 1. Technologies numériques et extension sociale de la cognition

Dans les études classiques de cognition distribuée (Norman, 1991 ; Hutchins, 1995a ; Kirsh, 1999, 2006 ; Hollan et al., 2000), l'extension sociale de la cognition est vue comme additive ou complémentaire à son extension technologique dans des outils et des artefacts. Une raison qui peut expliquer pourquoi les processus sociaux de distribution sont isolés des processus de coordination avec la technologie peut provenir de la classe d'artefacts qui était le plus souvent observée dans ces travaux. Ces études portent le plus souvent sur des artefacts de contrôle où la prise d'information est toujours plus ou moins liée à des coordinations directes entre un utilisateur et une interface. Le couplage agent/artefact se fait sous forme de routines d'exécution à base visuelle et motrice pour réaliser des tâches qui restent fortement contraintes par l'environnement. Dans un usage routinier, la contribution des pilotes à la mémorisation de la vitesse dans un cockpit (Hutchins, 1995b) est mineure par rapport à celle qui provient des indicateurs de vitesse, alors que la contribution des messages émis par les groupes d'expert dans une liste de diffusion est décisive pour acquérir des connaissances. Un artefact comme un indicateur de vitesse n'a, en effet, pas pour fonction principale de coordonner les pilotes ou de les faire communiquer car la dimension sociale de la distribution n'est pas inscrite dans la technologie<sup>3</sup>. Pour qu'un artefact intervienne directement sur la distribution sociale de la cognition, il faut qu'il agence la coordination homme–homme par exemple, en faisant communiquer les personnes. Dans ce cas, la distribution sociale n'est pas seulement complémentaire ou additive mais se loge au cœur de l'architecture de l'artefact. La question d'une élaboration de la connaissance par une communauté ne se pose donc pas si les artefacts de contrôle restent le vecteur principal de la distribution.

Une autre raison qui peut expliquer pourquoi l'interaction sociale et la communication ne se trouvent pas au centre de ces études viendrait de la relation qu'entretiennent les artefacts de contrôle avec l'acquisition de la connaissance. Les artefacts de contrôle, observés par D. Norman et E. Hutchins, ne visent pas à soutenir un *knowing usage*, c'est-à-dire des processus d'élaboration ou d'exploration de connaissance. Les artefacts cognitifs de contrôle soutiennent d'abord des processus d'exploitation de connaissances où les savoirs sont déjà acquis et implémentés dans des routines ou des heuristiques. La rationalité qui prédomine, en effet, dans un usage routinier des technologies numériques vise à diminuer le coût cognitif de la tâche en transformant, par exemple, des opérations de calcul en tâche à base visuelle et spatiale (Hutchins, 1995b ; Kirsh, 2006).

Cette relation entre artefact numérique et acquisition de la connaissance s'exprimerait d'abord à travers une modalité particulière de la coordination selon que la source de la connaissance provient principalement de l'artefact ou des personnes. Il semblerait qu'il existe une relation étroite entre les propriétés fonctionnelles non communicationnelles des artefacts de contrôle, la nature à base visuelle et spatiale des connaissances distribuées et le fait que la coordination se fait d'abord entre

<sup>3</sup> Les fonctions portant sur le contrôle de l'action qui interviennent sur des processus visuels-spatiaux et moteurs se distinguent des fonctions portant sur la construction des croyances qui interviennent sur la connaissance et la transmission des contenus. Les technologies informatiques évoluent en développant des systèmes qui interviennent de plus en plus sur les secondes fonctions et se transforment en *Mind Tools* au sens d'A. Clark (1998) et de D. Dennett (2000).

l'artefact et la personne qui l'utilise sous la forme d'une interaction homme-machine. Lorsqu'en revanche, la technologie numérique permet une augmentation de la communication entre les personnes, se présente un cas d'extension sociale de la cognition dans laquelle l'artefact facilite des coordinations homme-homme en les étendant et en accélérant l'accès à autrui. Les technologies numériques réticulaires agencent, en effet, des coordinations sociales sous forme de chaînes de communication qui deviennent le moteur principal de la dissémination des connaissances.

Il est donc logique que ce contraste opposant deux variétés de cognition distribuée corresponde à deux usages différents des artefacts cognitifs, un usage routinier dans le cas des artefacts numériques de contrôle et un usage épistémique dans le cas des artefacts numériques réticulaires. Une telle différenciation se manifeste dans le vocabulaire utilisé. Par exemple, les études qui portent sur la création de connaissances scientifiques ou sur l'exploration informatique dans la programmation en *Open-Source* parlent plutôt de « communauté épistémique » que de cognition distribuée. Ce terme exprime la dimension sociale de la distribution cognitive car il renvoie à des usages où la qualité des réseaux de communication et des échanges de connaissances joue un rôle de premier plan. Dans les usages épistémiques, l'organisation de la coordination sociale et l'intensité des interactions entre les agents sont beaucoup plus décisives que les *affordances* fonctionnelles d'artefacts qui agissent sur la coordination utilisateur-outil<sup>4</sup>. Doit-on pour autant réserver la notion de cognition distribuée à un usage routinier des artefacts de contrôle ? Plusieurs raisons semblent plaider pour ne pas réduire la cognition distribuée uniquement aux interactions homme-machine et aux usages routiniers des artefacts de contrôle et ainsi envisager son extension à des usages épistémiques.

D'abord, les artefacts cognitifs réticulaires liés à Internet et au Web 2.0 jouent aussi un rôle central d'aide cognitive externe qui facilite l'émergence des « communautés épistémiques » à travers les réseaux de communication électronique. Le Web 2.0 et Internet, en fournissant une panoplie d'outils pour enregistrer, stocker et diffuser des connaissances dans un langage public, jouent un rôle d'aides cognitives externes pour des usages épistémiques et assurent en même temps une base technologique pour coordonner des agents à distance au sein des communautés épistémiques en ligne.

Ensuite, la communication et l'utilisation d'un langage public mettent aussi en œuvre des processus de distribution de la cognition. Beaucoup de ce que nous apprenons provient des communications que nous avons avec autrui. Lorsqu'un informateur ~~me~~ transmet verbalement une connaissance, il joue un rôle d'aide et de complément cognitif équivalent à celui qu'apporte une technologie cognitive. Autrui agit alors comme un support externe, un pourvoyeur d'informations, sans qu'aucune technologie n'intervienne dans la distribution des informations (Clark et Chalmers, 1998)<sup>5</sup>.

Mais l'argument qui pèse le plus en faveur d'une extension de l'idée de cognition distribuée aux usages épistémiques réside dans le double effet introduit, d'une part, par les technologies réticulaires et, d'autre part, par la communication sociale. Lorsque les processus d'élaboration de connaissance prédominent sur les processus dédiés à son exploitation, les artefacts ne peuvent jouer pleinement leur rôle d'aide cognitive que dans la mesure où ils intensifient l'interaction sociale et élèvent en même temps le niveau de communication entre des agents qui sont à la

<sup>4</sup> Cela n'exclut pas qu'il existe des *affordances* sociales ou des indices qui facilitent la coopération dans les réseaux de communication électronique sous forme, par exemple, de commandes *reply-to* ou d'affichages des thèmes et des noms des partenaires dans les *headers* (en tête des fils de discussion).

<sup>5</sup> On peut considérer, cependant, que le langage agit dans la conversation comme un artefact (Clark, 1996). Le message d'autrui devient un artefact à travers les représentations publiques externes qu'il émet.

recherche des partenaires bien informés. Les agents, lorsqu'ils communiquent au sein d'un réseau électronique, ne cherchent pas principalement à routiniser leur connaissance, mais bien à acquérir des connaissances nouvelles et ils ne peuvent le faire sans communiquer intensément entre eux. Lorsque les usages épistémiques prédominent, le but que se donnent les agents n'est plus de simplifier des connaissances maîtrisées au moyen d'ustensiles informationnels, mais d'explorer des connaissances évolutives non encore maîtrisées (Conein, 2004c, 2006). Ce qui distingue les artefacts réticulaires liés au développement d'Internet comme les listes de diffusion des artefacts de contrôle observés par D. Norman (1991) et par E. Hutchins (1995b), c'est le fait que ces artefacts coordonnent un nombre important de personnes pour communiquer des conseils et des idées nouvelles.

Si l'importance de la coordination sociale et de la communication a donc été sous-estimée dans certaines études sur la cognition distribuée, ce serait d'abord parce que celles-ci se sont penchées sur des classes de technologies numériques qui n'étaient pas pourvues de fonction communicative. La cognition distribuée n'est pas qu'une hypothèse sur l'interaction homme-machine, elle n'est pas limitée à ce type de paradigme technologique car elle concerne tous les mécanismes cognitifs susceptibles d'étendre les capacités cognitives au-delà des limites d'un organisme naturel (Conein, 2004c). L'observation des échanges de connaissance par les utilisateurs *Debian* exprime donc d'abord un usage épistémique des artefacts cognitifs où ce sont des réseaux plus ou moins denses d'agents en interaction qui acquièrent et produisent de la connaissance. Elle se distingue ainsi de l'observation des interactions homme-artefact où ce sont des usages routiniers des technologies numériques qui restent la source principale de l'aide cognitive<sup>6</sup>. Si l'extension sociale de la cognition n'est plus seulement, dans ces contextes, complémentaire à son extension au moyen des artefacts, c'est parce que les nouveaux artefacts cognitifs favorisent directement l'extension sociale de la cognition.

## 2. Fils de discussion et structures de répliques

Les analyses présentées portent sur le trafic, observé à partir d'août 1997, dans une liste de diffusion des utilisateurs francophones de la distribution du logiciel libre : *Debian-french*. Sont traitées au sein du même corpus, à partir de 2001, les deux listes, celle principalement destinée aux utilisateurs (*Debian-user-french*) et celle principalement destinée aux développeurs (*Debian-devel-french*)<sup>7</sup>. Les données considérées ici rassemblent l'ensemble des messages postés sur la *mailing-list* dès sa création (date à date). Elles ont été obtenues grâce à un programme spécialisé *Python*, développé pour récupérer dix ans de trafic et dont la fonction a été de collecter tous les messages à partir des archives disponibles en ligne.

L'objectif de l'enquête est d'exploiter les ressources de la théorie des graphes à traiter des grands corpus pour saisir des dynamiques relationnelles. Elle permet de voir émerger des structures

<sup>6</sup> Ces remarques doivent être, cependant, modérées de deux façons. D'abord, les utilisateurs du logiciel libre travaillent sur un ordinateur personnel et sont donc engagés dans une coordination homme-ordinateur à travers une interface de contrôle. Ensuite, un usage épistémique suppose la mise en œuvre des supports externes qui facilitent le raisonnement, l'évaluation et la planification en fournissant des aides visuelles (voir Kirsh, 2005).

<sup>7</sup> *Debian* est une distribution du logiciel libre qui est né en 1993 grâce à deux étudiants de l'université de Purdue : Deborah et Ian Murdock. *Debian* a reçu le soutien du projet GNU de la *Free Software Foundation*. La distribution *Debian* existe selon trois versions : une version stable, une version *test* et une version instable. En 2001, la *mailing-list Debian-french* se sépare en deux listes : une liste principale dédiée à l'usage des logiciels libres et une liste regroupant les développeurs. Dans la liste *user*, tous les membres, utilisateurs comme développeurs, se retrouvent.



relationnelles à partir des dynamiques d'interaction sur des périodes longues. L'avantage de la méthode est d'éviter d'étiqueter les agents avec des rôles et les relations avec des contenus avant l'examen et le traitement de propriétés relationnelles simples. Dans le cas d'une *mailing-list*, les propriétés relationnelles simples sont : qui réplique à qui ? Qui répond à qui ? Ce n'est qu'en un second temps que l'on regarde si ces propriétés peuvent générer d'autres caractéristiques plus complexes, comme par exemple : qui discute avec qui ? Qui conseille qui ? Bien que l'objet central de l'enquête soit d'identifier des relations de discussion, celles-ci sont conçues comme le produit de relations plus simples (comme des relations de répliques). De la même façon, les rôles de discutant ou de conseiller pourraient se construire à partir des relations de répliques.

Un second aspect de l'enquête concerne la relation entre diffusion de la connaissance et élaboration de la connaissance. Cette distinction peut-elle être rendue manifeste en observant les *patterns* de répliques ? Par exemple, si les messages de répliques dans un fil de discussion sont principalement des réponses, peut-on dire que la fonction de diffusion de la connaissance devient prédominante ? Si, en revanche, les répliques sont des répliques à des réponses, peut-on dire que la connaissance attachée au message tend à être objet d'un processus de transformation élaborative ?

Dans une liste de diffusion, les réseaux de communication électronique agencent, en effet, les relations de communication entre émetteur et destinataire sous forme d'un fil de discussion (*thread*) organisé autour d'un thème. Chaque fil de discussion est ouvert par un message de requête postée (*ml*) qui est suivie d'une ou plusieurs répliques, mais peut aussi ne pas l'être.

Dans *Debian-user-french*, la requête prend souvent la forme d'une demande de conseil en matière de programmation, postée sur la liste, exprimée par le titre de l'en-tête et dans le contenu de la question et de la réponse :

Aptitude autoremove

A1 : « Bonjour, dans la doc j'ai vu que *aptitude* était capable de supprimer les *deb* installés automatiquement après un *apt-get install xxx* après testé différentes options (*-purge-unused*...) quelqu'un pourrait-il me donner la bonne ligne de commande pour faire cel, merci »

B1 : « *aptitude install deborphan* ça peut t'aider »

Pour caractériser une structure de répliques, il faut tenir compte des relations entre les messages au sein de chaque fil de discussion. Chaque fil de discussion peut générer, en effet, deux classes de répliques<sup>8</sup> : des *relations de réponse* lorsque la réplique répond à la question posée dans la requête postée et des *relations de réplique qui ne sont plus des réponses* lorsque les répliques s'adressent à une réplique antérieure<sup>9</sup>. Ce phénomène a été observé par Welser et al. (2007) qui soulignait que les fils de discussion ont tendance à générer d'autres relations de répliques que des relations question-réponse.

Si on représente les fils de discussion par des graphes, on peut donc mettre à jour deux sous-structures de répliques entre les messages (Fig. 1).

Dans le premier cas, les répliques s'orientent vis-à-vis de la question, le questionneur est sélectionné comme destinataire et l'auteur du fil se maintient à un. Dans le second cas, elles s'orientent vis-à-vis d'une réplique antécédente ; le questionneur n'est plus le destinataire principal et l'auteur du fil augmente. Un fil de discussion peut donc se représenter selon deux graphes de répliques, dont la structure est clairement distincte. Lorsque le fil se présente comme un *réseau*

<sup>8</sup> L'orientation des répliques vers le questionneur ou vers le répondant est exprimée de deux manières : soit dans le message, par la citation du message à qui on répond, soit dans l'en-tête par l'adresse électronique citée.

<sup>9</sup> Si une requête postée est presque toujours une question qui invite à une réponse dans le message suivant, la réplique à une réponse est plus une évaluation de la réponse qu'une réponse à la requête.

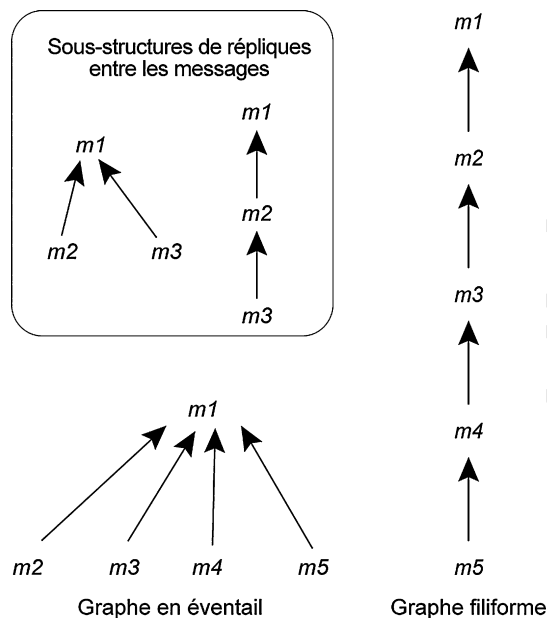


Fig. 1. Réseau de réponses (graphe en éventail) et réseau de répliques (graphe filiforme) dans les fils de discussion.

de réponses, le graphe tend à prendre une forme en éventail. Lorsque le fil se présente comme un réseau de répliques, le graphe prend une forme filiforme (Dorat, 2005), puisque chaque réplique s'oriente vers la réplique précédente.

Cette dualité de forme dans les structures de répliques va nous permettre de mettre en valeur trois phénomènes peu encore pris en compte : l'intensification des interactions sociales, l'augmentation de la communication et la transformation de la connaissance.

### 3. Réseaux de communication et intensification des interactions sociales

Dans la communication, pour acquérir une connaissance, le procédé le plus direct est, en effet, qu'une personne pose une question d'information à un informateur. Ce n'est donc nullement un hasard si les standards de *requests for comments* et les protocoles *Usenet*<sup>10</sup> utilisent des couplages question/réponse pour transmettre des messages.

En enregistrant, stockant et diffusant des flux des questions et des réponses à une liste d'individus, les listes de diffusion (*newsgroups*) permettent de suivre, sur des périodes longues, les interactions entre questionneurs et informateurs dans les fils de discussion<sup>11</sup>. Or la forme des fils

<sup>10</sup> Les systèmes *Requests for comments* (RFC) sont utilisés par les informaticiens depuis 1969 et aujourd'hui par les usagers d'Internet. Les protocoles RFC agencent des standards de communication et archivent des documents. *Usenet* est un ensemble de protocoles pour les forums de discussion et les listes de diffusion. Les membres de la liste peuvent recevoir les messages et les articles diffusés et ceux qui sont archivés. La liste *Debian.user:french* utilise comme la plupart des listes de diffusion les protocoles Usenet.

<sup>11</sup> L'idée de traiter les fils de discussion comme des réseaux doit beaucoup aux séances de travail avec R. Dorat (LIFL) et M. Latapy (LIP6). Sans les projets programme d'étude des réseaux sociaux de l'Internet (PERSI) et l'ANR « Autograph », cette réflexion sur les réseaux de communication n'aurait pas pu être développée.

de discussion présente des propriétés relationnelles qui ont des effets sur l'acquisition et la création de connaissances (Latapy et al., 2007). L'observation<sup>12</sup> de la dynamique des échanges, dans les *threads*, entre les utilisateurs de *Debian* se révèle alors une source précieuse d'information, en particulier concernant le rôle des interactions sociales comme support aux relations de discussion.

Un premier fait intéressant porte sur la façon dont la dynamique des interactions contribue à l'extension sociale de la cognition. Une façon de comprendre l'extension sociale de la cognition serait en effet de montrer que les interactions sociales ont des propriétés émergentes. Par exemple, l'augmentation du nombre d'interaction entre les contributeurs de la liste pourrait avoir des effets sur l'apparition de relations dialogiques de discussion. L'établissement de ce lien n'est pas évident au départ. Rien ne permet de dire qu'une élévation du niveau de connaissance résulte toujours d'une intensification des interactions entre les agents. Cependant, dans le contexte de la liste des utilisateurs *Debian* qui rassemble des compétences techniques de haut niveau dans la programmation informatique, l'augmentation des interactions entre les répondants peut être considérée comme une des expressions du niveau de connaissance des contributeurs.

Dans ce contexte, un niveau élevé d'interaction entre les agents pourrait s'avérer un facteur favorisant la communication des idées et l'émergence de processus de discussion orientés vers l'élaboration de la connaissance. Le nombre de messages postés par un auteur est généralement conçu comme un bon indice de l'augmentation du taux de connaissances dans la liste de diffusion (Lakhani et von Hippel, 2000). Mais ce n'est pas le seul indice. L'augmentation du taux d'interaction sociale peut s'avérer un meilleur indice de l'augmentation du taux de connaissance car il crée des conditions favorables à l'émergence de relations dyadiques de discussion.

Un tel phénomène est facilement observable lorsqu'on compare le nombre de messages dans un fil de discussion au nombre d'auteurs de messages (Fig. 2). Le taux d'interaction étant mesuré par la relation nombre de messages/nombre d'auteurs, ce taux peut être alors conçu comme l'expression d'une réciprocité dans la relation entre contributeurs au sein d'un fil de diffusion.

Pour que des interactions mutuelles répétées émergent, les auteurs de message doivent pouvoir intervenir une seconde fois, voire plusieurs fois. Ce fait se traduit par une tendance à ce que le nombre d'individus, auteurs d'un message, tend, par conséquence, à stagner ou à s'accroître d'une manière plus lente au fur et à mesure que le nombre de messages augmente dans un fil de discussion.

Un second fait s'avère plus intéressant pour l'étude des relations de discussion et des usages épistémiques des artefacts numériques. Le nombre d'interactions répétées augmente selon la forme que prend le graphe des répliques dans un fil de discussion. La forme de ce graphe, nous venons de le souligner, évolue entre deux cas typiques : une forme en éventail ou une forme filiforme (Latapy et al., 2007).

Si ce second fait s'avère plus intéressant pour notre analyse, c'est qu'il souligne bien que le taux d'interaction et donc la redondance des auteurs dans le fil, va augmenter de façon beaucoup plus manifeste lorsque les relations entre auteurs/messages s'expriment dans une structure de graphe filiforme (Fig. 3).

Le caractère filiforme de tout fil est mesuré par sa hauteur, c'est-à-dire par la longueur de la plus longue suite de répliques contenues dans le fil. Cela permet de prendre en compte tous les messages et non plus seulement les fils strictement filiformes, comme dans la Fig. 3. La tendance des fils les plus filiformes à contenir plus d'interactions est alors confirmée (Fig. 4).

<sup>12</sup> Deux observations ont été menées : d'abord, sur un corpus de 25 941 messages postés en 2003 à 2004, puis sur un corpus de 158 383 messages en 1997 à 2007.



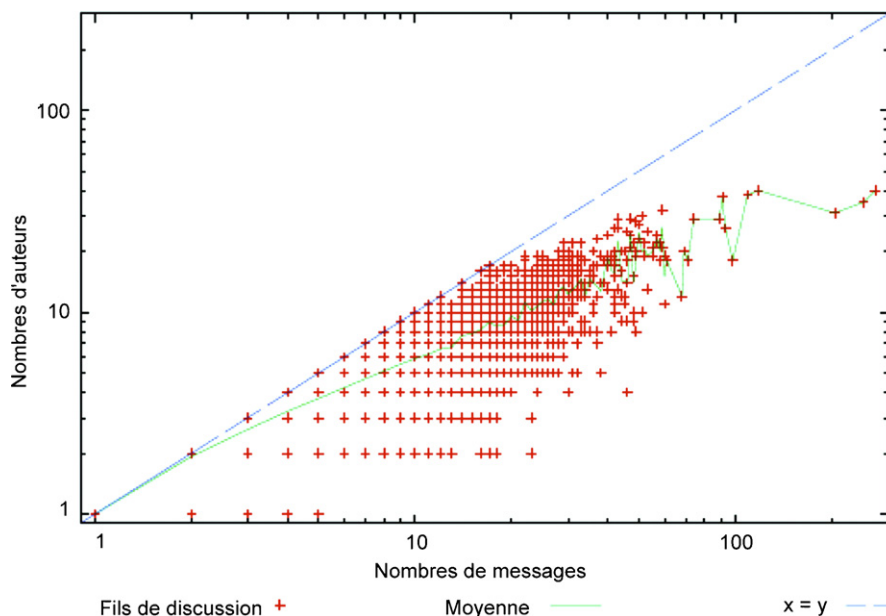


Fig. 2. Pour chaque fil de discussion, son nombre d'auteurs en fonction de son nombre de messages dans la liste *debian.user.french* (1997–2007). Dans ce graphique, chaque point correspond à un fil de discussion, ses coordonnées  $(x, y)$  indiquant son nombre de messages  $x$  et son nombre d'auteurs  $y$ . La moyenne de ces valeurs, c'est-à-dire le nombre moyen d'auteurs, pour chaque valeur rencontrée du nombre de messages, est exprimée par la ligne brisée. Notons que le nombre d'auteurs ne peut pas être supérieur au nombre de messages, donc tous les points sont en dessous la droite diagonale  $y = x$ . Afin de rendre la courbe lisible, malgré l'hétérogénéité des valeurs, les axes sont logarithmiques. Il apparaît clairement que, plus le fil contient des messages, plus son nombre d'auteurs a tendance à être comparativement faible (les points s'éloignent de la diagonale).

Par contraste avec l'appréciation de l'hauteur d'un fil, on peut estimer le caractère en éventail d'un fil par sa largeur, mesurée par le degré, c'est-à-dire le nombre maximum de répliques directes engendrées par un message dans le fil. Il apparaît alors clairement (Fig. 5) que les fils les plus larges contiennent moins d'interactions.

Il apparaît clairement (Fig. 5) que le taux d'interaction a tendance à cesser de croître quand le degré croît (les derniers points ne sont pas significatifs car représentent très peu de fils). Cela indique que le caractère en éventail d'un fil révèle une faible redondance des auteurs dans ce fil, qui révèle à son tour une faiblesse des interactions. Le taux d'interaction augmente beaucoup plus faiblement (voire se stabilise, Fig. 5) en fonction de la largeur qu'en fonction de la hauteur (Fig. 4). Cela montre que les fils larges contiennent moins d'interactions, comparativement, que les fils hauts (bien que persiste la tendance des fils contenant le plus de messages à contenir plus d'interactions).

Quand le fil tend donc à prendre une forme filiforme (que donc sa hauteur augmente) et lorsque le nombre de messages aussi augmente, la stagnation dans l'accroissement du nombre d'auteurs devient plus manifeste. L'apparition d'interactions répétées lorsqu'elle se fait conjointement avec une intensification des relations de corépliques permet de mieux voir comment des sous-structures de discussion peuvent émerger à l'intérieur des structures de communication initiées par une question d'information.

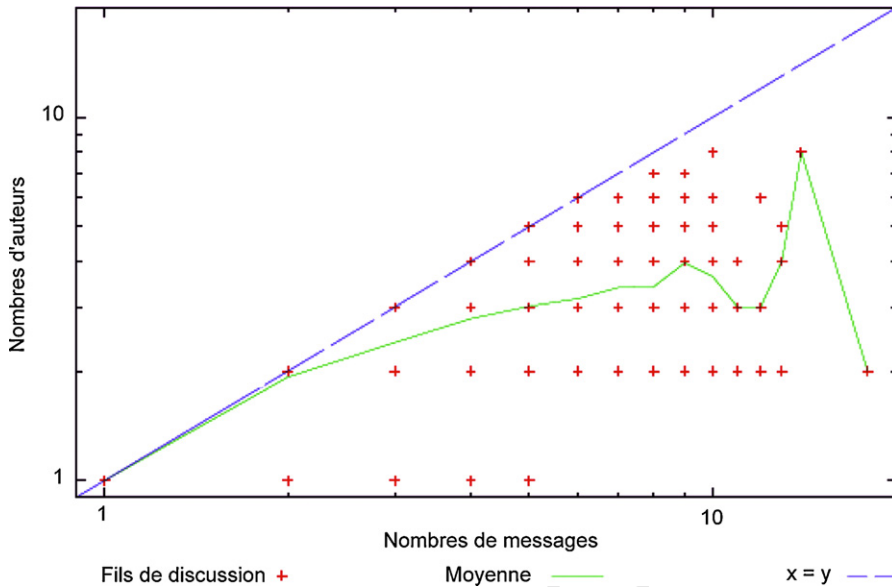


Fig. 3. Pour chaque fil de discussion filiforme, son nombre d'auteurs en fonction de son nombre de messages dans la liste *debian.user.french* (1997–2007). Cette figure est la même que la Fig. 2, mais cette fois-ci on ne considère que les fils de discussion purement filiformes. On constate ainsi que la tendance à s'éloigner de la diagonale, lorsque le nombre de messages croît, semble plus forte pour les fils filiformes que dans le cas général. Toutefois, le faible nombre de fils strictement filiformes rend limitée la portée des inférences à tirer de cette courbe.

Les deux phénomènes significatifs sont donc la forme du graphe du fil de discussion (sa hauteur et sa largeur) et la distinction entre réponses et répliques. Tous les messages qui ne sont pas des questions ne sont pas automatiquement des réponses. L'usage épistémique des listes reposerait alors sur la possibilité que se développent des relations de réplique qui ne sont pas des réponses au sein des fils de discussion.

#### 4. Réseaux de communication et augmentation de la communication

Le contraste entre des réseaux où prédominent des relations de réponses et des réseaux où prédominent d'autres types de répliques concerne d'abord la fluidité de la transmission et de la circulation des messages au sein d'une liste de diffusion. Cela peut vouloir dire que les technologies réticulaires ne sont pas seulement des systèmes d'extension de la diffusion. Augmenter la communication entre les membres de la liste ne se réduit pas à disséminer de l'information, mais impliquerait des formes de transmission plus faiblement préservatives du contenu transmis car plus fortement productives de connaissances nouvelles (Sperber et Claidière, 2006). Augmenter la communication exprimerait ainsi deux phénomènes distincts : augmenter la diffusion en augmentant le nombre de destinataires et augmenter la production en augmentant le nombre de contenus distincts.

À première vue, l'augmentation de la communication dans un réseau numérique comme celui de *Debian.user.french* porte d'abord sur la diffusion d'informations spécialisées de qualité pour les abonnés de la liste ; c'est-à-dire à l'intention d'un nombre important de destinataires. Les *news groups* ont d'abord été conçus comme des outils de diffusion de *news* et d'*articles*. La fonction

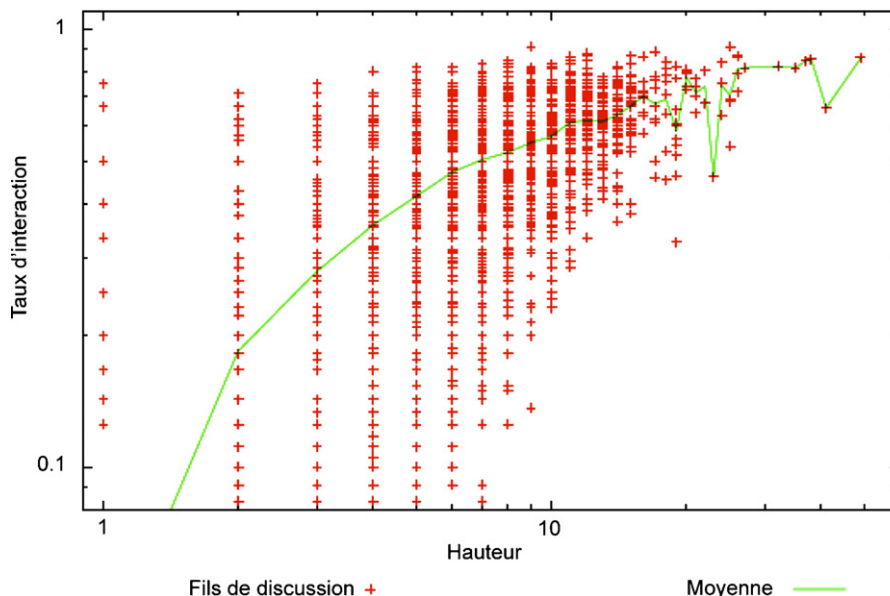


Fig. 4. Taux d'interaction en fonction de l'hauteur (mesure du caractère filiforme du fil de discussion) dans la liste *debian.user.french* (1997–2007). Chaque point correspond à un fil de discussion, ses coordonnées (x, y) indiquant sa hauteur x et son taux d'interaction y. La moyenne de ces valeurs, c'est-à-dire le taux d'interaction pour chaque valeur rencontrée de l'hauteur des fils est représentée par une ligne brisée. Afin de rendre compte au mieux de l'hétérogénéité des valeurs rencontrées, les axes sont en échelle logarithmique. Il apparaît clairement que le taux d'interaction croît très vite (tant la moyenne que les valeurs minimum et maximum) en fonction de l'hauteur, indiquant que le caractère filiforme d'un fil est induit par une forte redondance des auteurs dans ce fil, qui indique à son tour une forte interaction dans le fil.

de diffusion est bien facilitée par la numérisation des réseaux de communication qui produit une accélération de la dissémination des messages pour un nombre élevé de destinataires. De plus, l'organisation des listes de diffusion par émission de requêtes postées oriente la production des messages vers une production de réponses. La nature de la connaissance semble donc indirectement être affectée par la modalité de sa transmission. Or la possibilité de donner une réponse à une question suppose que la connaissance transmise est standardisée sous forme d'une instruction et n'est pas transformée à chaque transmission. Ce qui importe, c'est donc le nombre de destinataires qui peuvent utiliser cette instruction pour programmer. Le premier objectif d'une liste de diffusion comme celle de *Debian* est donc de disséminer des connaissances « codifiées » aux membres de la liste qui cherchent assistance et conseil. Ce qui explique que dans le cas du fil déjà cité (aptitude *autoremove*), la réponse à la requête se réduit à une ligne de code : « *aptitude install deborphan ça peut t'aider* » et que le second conseil prendra aussi la forme d'une instruction alternative : « ou même utiliser *orphaner-purge* »<sup>13</sup>.

Mais si la fonction de diffusion est réalisée par des typages des messages selon des formats question-réponse, elle a besoin d'accueillir aussi des messages de répliques qui ne sont pas des réponses, mais des critiques, des évaluations ou des commentaires. Dans le même fil, le dernier message prend la forme de l'explicitation de l'usage d'une procédure : « Aptitude peut (et normalement le fait) désinstaller les programmes qui sont marqués comme "auto" (il y a un

<sup>13</sup> Cf. archive de la *Debian-user-french mailing list* : fil *aptitude autoremove*, requête postée le 18 mars 2008.

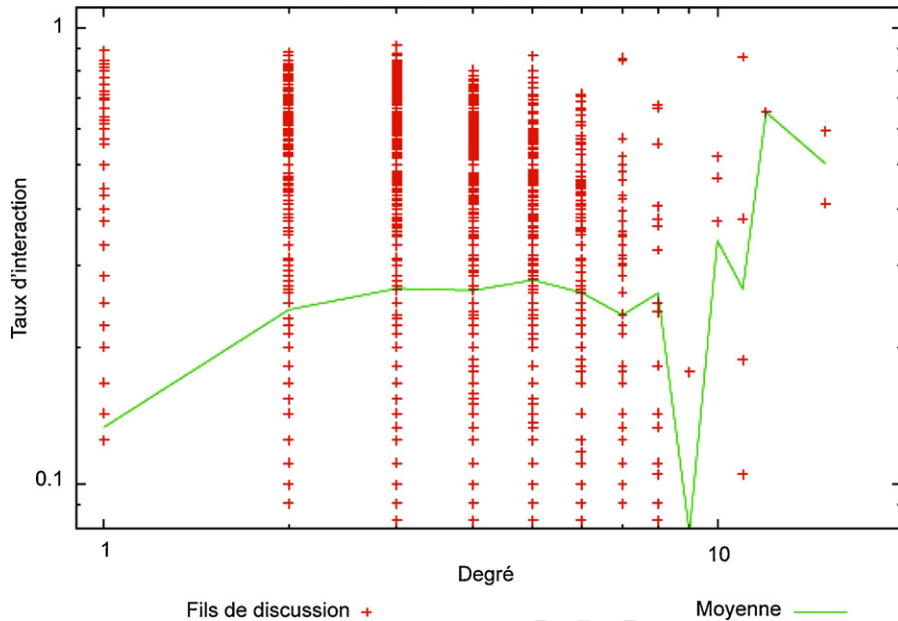


Fig. 5. Taux d'interaction en fonction de la largeur ou du degré (mesure du caractère en éventail du fil de discussion) dans la liste *debian.user.french* (1997-2007). Chaque point correspond à un fil de discussion, ses coordonnées (x, y) indiquant son degré x et son taux d'interaction y. La moyenne de ces valeurs, c'est-à-dire le taux d'interaction pour chaque valeur rencontrée du degré des fils est représentée par une ligne brisée. Afin de rendre compte au mieux de l'hétérogénéité des valeurs rencontrées, les axes sont en échelle logarithmique. De même, pour rendre la courbe plus lisible, nous avons commencé l'axe des y à 0,08, valeur minimale atteinte par la moyenne. Des valeurs inférieures existent toutefois (non représentées) qui sont responsables de la chute de la moyenne en  $x=9$  (elle serait sinon aberrante).

grand A à côté du nom) quand il y a plus rien qui dépend d'eux. Si tu as installé quelque chose à la main, même si c'est inutile, il va rester (sauf si tu fais un *remove* à la main) ».

Une part importante des échanges conduit à mettre en cause le mode conceptuel de celui qui poste la requête et à lui faire acquérir une compétence nouvelle, alors que dans d'autre cas, le questionneur peut se contenter de copier le code qu'on lui propose pour appliquer la procédure<sup>14</sup>. Dans le premier cas, les connaissances attachées aux messages deviennent évolutives car elles font l'objet d'un processus de transformation et donc de révision partielle. En conséquence, la fonction de diffusion devient moins centrale, puisque l'acquisition de connaissance devient véritablement créative. Ce qui a pour conséquence que l'augmentation de la connaissance au sein de la communauté se mesure en termes de transformation de contenus et plus seulement en termes de nombre de destinataires.

A. Goldman souligne que l'augmentation de la communication dans l'acquisition de connaissance présentera alors deux modalités assez distinctes, une faiblement coopérative car principalement orientée vers la transmission d'un contenu stable et une collaborative orientée vers l'élaboration mutuelle de nouveaux contenus :

<sup>14</sup> Nous devons cette remarque sur l'usage de la connaissance informatique à R. Dorat. Le couplage entre *skilled knowledge* et connaissance formelle serait une des caractéristiques de cette connaissance.

« La communication peut jouer un rôle critique à la fois dans la dissémination de connaissances anciennes et dans l'acquisition de connaissances nouvelles. Dans le premier cas, la communication est requise pour transmettre des messages porteurs de vérité. Dans le second cas, où la collaboration est concernée, la communication entre collaborateurs devient essentielle. » (Goldman, 1999, p. 161).

A. Goldman relie deux modalités de transmission à des modifications qui portent sur les contenus et les états de connaissance. Dans le cas d'une diffusion d'une connaissance standardisée, la fonction de la transmission serait principalement préservative du contenu du message. Une réponse à une question vise d'abord à transmettre ce qui est connu. Ce qui conduit à dire que les fils où prédominent des répliques qui sont des réponses visent d'abord à diffuser des connaissances au sein de la liste et donc à augmenter la connaissance selon le premier sens.

En revanche, dans le cas de l'acquisition d'une connaissance évolutive, le processus de transmission serait de nature élaborative ou constructive, puisque la fonction ne vise plus à préserver un contenu ou à sélectionner une signification, mais à les transformer en générant un nouveau contenu (Sperber et Claidière, 2006)<sup>15</sup>. Les répliques qui ne sont pas des réponses tendent à initier un processus de coconstruction de connaissance.

Pour M. Tomasello, ces distinctions dans les modes de transmission ont des conséquences pour l'apprentissage (Tomasello, 1993, 1998). Le premier cas correspondrait à un apprentissage de type instructionnel où la qualité de la transmission est centrale pour que la compréhension du novice soit la meilleure possible. Le lien entre apprentissage et nature des connaissances va se traduire par le fait suivant : la réponse à la question d'un novice va tendre à prendre une forme instructionnelle et « codifiée ».

Au contraire, le second cas correspond à un apprentissage collaboratif<sup>16</sup> où la reconstruction du contenu est centrale pour engendrer des contenus nouveaux d'une manière créative. Ainsi, un intervenant chevronné est plus intéressé à acquérir des connaissances nouvelles sur un problème fondamental que des instructions. Un novice peut aussi être intéressé à connaître les principes d'arrière-plan sous-entendus par l'acquisition de la procédure. La dimension collaborative et discutante semble directement liée à la possibilité de communiquer entre pairs, soit sur des problèmes fondamentaux d'architecture, soit sur les principes-cadre qui sont derrière les procédures ou les instructions<sup>17</sup>. Il existe donc des cas intermédiaires où la transmission d'une instruction est reliée à l'exposition d'un principe qui permet au novice d'adapter la réponse à son cas, mais aussi de la reproduire dans d'autres cas.

La distinction entre ces deux modes de transmission permet donc de ne pas réduire l'augmentation de la communication à un simple phénomène de diffusion. Sur ce point également, la distinction entre réponse et réplique s'avère pertinente.

<sup>15</sup> D. Sperber et N. Claidière proposent une autre distinction qui complète celle d'A. Goldman : « toute transmission d'information subit une double contrainte : une contrainte de préservation du contenu et une contrainte de transformation » (Sperber et Claidière, 2006).

<sup>16</sup> Dans le cas de l'apprentissage collaboratif, « l'apprenti coconstruit avec un pair, ayant les mêmes connaissances et un statut équivalent, une nouvelle structure cognitive, il s'agit d'un processus de création culturelle ou de construction plutôt que de transmission » (Tomasello et Ratner, 1993).

<sup>17</sup> R. Dorat souligne (communication personnelle) que l'opposition entre les connaissances instructionnelles basées sur des aptitudes (*skills*) et les connaissances évolutives non standardisées est relative dans le cas de la connaissance informatiques. La relation entre *skilled-knowledge* et connaissance au sens strict serait très forte car la démarche de programmation implique toujours des procédures qu'on doit suivre.

Si on revient sur la place occupée par la question dans la communication des idées, on peut soutenir que le couple question/réponse assure principalement une fonction préservatrice de diffusion car l'informateur pour conseiller doit fournir une réponse standard qui doit en même temps être ajustée au niveau de compréhension du questionneur et des membres novices sur la liste. L'informateur ne cherche pas, dans ce contexte, à élaborer une connaissance ou modifier le modèle conceptuel du novice, mais à transmettre ce qui est déjà connu pour une application, soit une connaissance ancienne relativement stabilisée. Inversement, le couple question/réponse semble moins adapté à une communication « discutante » où l'échange d'idées nouvelles, le renvoi à des principes ou des problèmes fondamentaux devient le but de ceux qui participent à la discussion. Bien qu'une discussion commence souvent par une question, la mise en discussion d'une réponse est un des chemins commodes pour initier un renvoi à des questions de fond au sein d'un *thread*. La mise en discussion des réponses altère donc le jeu de langage question/réponse en instaurant une dynamique collaborative dans les relations de réplique.

La communication numérique s'avèrerait un instrument efficace d'augmentation de la connaissance dans la mesure où elle facilite les deux manières d'augmenter la communication au sein d'une communauté unique. Lorsqu'un intervenant sur une liste de discussion a l'intention, ou accepte, d'informer, en proposant un conseil ou une idée, il devient un support pour la diffusion de la connaissance (Davis, 2002). Les listes, en facilitant en même temps la diffusion des conseils à un public élargi et la construction d'idées, pallient une faiblesse de la communication verbale qui ne peut rassembler qu'un nombre réduit à la fois de destinataires et d'informateurs. Les contraintes de coprésence et d'attention conjointe, deux caractéristiques selon E. Goffman de la conversation (Goffman, 1961), limitent autant la diffusion que la discussion des idées. Un des modes particuliers d'intégration de la conversation dans les *requests for comments* sur les listes de diffusion consistera justement à préserver certaines propriétés interactives de la relation question/réponse tout en surmontant les limites qu'introduit la contrainte de coprésence pour la diffusion en réduisant le nombre des participants et d'interactions entre eux<sup>18</sup>.

## 5. Réseaux de communication et transformation de la connaissance

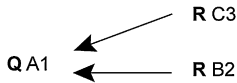
Comment ces modifications de la forme des réseaux de communication affectent-elles les réseaux de connaissance au sein de la liste ? Si la forme des fils de discussion exerce bien des contraintes sur la communication des messages et leur mise en séquence, elles pourraient également s'exercer sur la transformation des connaissances et le mode d'interaction entre les informateurs.

Les propriétés qui caractérisent un réseau de connaissance ne correspondent pas à celles d'un réseau de communication. Par exemple, si les notions de discussion comme de conseil concernent la connaissance, les notions de transmission et de répliques concernent la communication. Un réseau de communication reste un réseau qui met en relation des messages orientés, c'est-à-dire des répliques. Les deux types de réseau ne sont pas équivalents car leurs graphes peuvent être différents. Si on peut exprimer la forme d'un réseau de communication numérique par l'orientation des répliques, en même temps, il n'y a pas de corrélation systématique entre graphe en éventail et

<sup>18</sup> Un débat s'est instauré dans les études de communication médiée par ordinateur : savoir s'il s'agit encore de conversation. De nouveaux termes sont apparus comme *threaded conversation* ou *persistent conversation* pour désigner les propriétés conversationnelles de la communication électronique en tant que mode de communication hybride. M. Marcoccia critique, avec de bons arguments, cet emploi de la notion de conversation pour analyser la communication électronique (Marcoccia, 2004).



**Configuration 1 : trois agents, deux réponses (graphe en éventail)**



**Configuration 2 : trois agents, une réponse et une réplique évaluative (graphe filiforme)**



**Configuration 3 : deux agents, une réponse et une réplique évaluative du questionneur (graphe filiforme)**



Fig. 6. Configurations relationnelles minimales au sein d'un fil de discussion.

relation de conseil et entre graphe filiforme et relation de discussion. En revanche, il existe bien une réelle correspondance entre forme large en éventail et cycle question/réponse et entre forme haute filiforme et cycle réplique à une réplique antécédente.

Ce qui veut dire que lorsque le nombre de messages dans un fil de discussion augmente, l'acquisition de connaissances semble dépendre de deux cycles interactifs distincts : des cycles *question/réponses* et des cycles *question/réponse/répliques évaluatives*. La relation de discussion tire parti d'une propriété que comporte le second cycle : une réplique peut évaluer une réponse et ouvrir la possibilité que le répondant réplique à son tour ~~et donc s'aligner mutuellement~~. La différenciation des répliques entre réponse et réplique évaluative, dans ce cadre, devient fondamentale car le second cycle ouvre à des interactions réciproques à la différence du premier. C'est cette réciprocité fondée sur un alignement épistémique (Lazega, 2001) entre les répliquants qui est susceptible de créer des conditions favorables pour des relations de discussion. Certains cycles interactifs dans les fils de discussion seraient alors plus favorables que d'autres, soit pour diffuser des connaissances stables au plus grand nombre sous forme de conseils, soit pour transformer des connaissances transmises à l'intérieur des niches de compétence. Par exemple, la forme en éventail d'un fil de discussion serait limitative du point de vue informationnel car les intervenants ne considèrent que la question et n'interagissent qu'avec le questionneur, alors que la forme filiforme est ouverte non seulement parce qu'elle permet d'ajouter une information nouvelle à chaque réplique, mais parce qu'elle initie les interactions répétées et réciproques.

Un moyen commode d'estimer la transformation de la connaissance serait de concevoir l'impact des réseaux de communication sur l'apparition de relations de discussion comme des filtres où les messages peuvent être couplés de plusieurs façons. Une façon de le faire serait de considérer des configurations<sup>19</sup> relationnelles minimales, afin de voir à quel point un fil de discussion peut être interactif, collaboratif ou critique selon l'orientation des répliques (Fig. 6).

Les *patterns* de réplique peuvent, en effet, être représentés sous forme de trois graphes triadiques couplant des messages orientés (Q/R/R') et des auteurs (A, B, C) occupant des rôles différents (questionneur, informateur/répondant, informateur/discutant)<sup>20</sup>. Pour trois messages émis par A, B, C, on peut savoir si B et C construisent des réponses à A ou si B produit une

<sup>19</sup> En analyse de réseaux, la notion technique de configuration est une schématisation. Elle forme un sous-ensemble de nœuds et de relations (arcs) qui peut être contenu dans une triade.

<sup>20</sup> On peut en fait construire un graphe dont les nœuds sont des messages, puis un graphe dont les nœuds sont des agents, soit un graphe de messages étiquetés (cf. Dorat et al., 2006).

réponse à A et C une réplique à la réponse de B. Chaque agent est alors assigné à une position particulière selon sa réplique et sa relation à un autre agent.

Dans la première configuration Q/R/R, le graphe est de largeur 2 (ce qui est sa largeur maximale pour deux répliques), le taux d'interaction est faible dans un rapport auteurs/messages de 3/3 (configuration 1, Fig. 6). En même temps, chaque réplique prend la forme de deux messages de réponses à une question, ce qui favorise à la fois un processus de diffusion et une relation de conseils car le questionneur A, mais aussi les abonnés de la liste peuvent comparer la convergence des deux réponses, celle donnée par B et celle ensuite donnée par C<sup>21</sup>.

Dans la seconde configuration Q/R/R', le graphe est de largeur 1 (ce qui est la largeur minimale pour deux répliques) mais le taux d'interaction ne change pas (3/3) (configuration 2, Fig. 6). Les deux répliques bifurquent car si B répond à A, C réplique au message de B. La production d'une nouvelle réplique rend possible une interaction collaborative critique et donc une transmission élaborative entre les deux informateurs B et C. Le questionneur, mais aussi les abonnés de la liste, peuvent non seulement comparer les deux types de répliques, mais tirer bénéfice des informations nouvelles apportées par la discussion entre les deux informateurs B et C.

La dernière configuration ressemble à la seconde, mais avec une modification du taux d'interaction (configuration 2, Fig. 6). Si le graphe est de largeur 1 (ce qui est sa largeur minimale pour deux répliques), le taux d'interaction change car A réintervient. L'orientation vers la discussion est accentuée car elle introduit une forme nouvelle de distribution de la connaissance entre les intervenants : A en réintervenant évalue la réponse de son informateur. Elle présente un cas particulier d'insertion d'une structure collaborative de discussion au sein de la demande de conseil en modifiant à la fois l'orientation de la réplique et le taux d'interaction (rapport de 2/3).

Ce cas de figure peut exprimer la production d'une question élaborative de type construction de problèmes et engendre une structure collaborative avec interaction mutuelle à deux auteurs. À change de rôle communicationnel et évalue le complément de connaissance apporté par la réponse de son informateur, qui devient ainsi son collaborateur.

Dans ces trois exemples, on voit que l'on peut estimer en termes de distribution sociale de la connaissance, à la fois la communication sous forme de dissémination des messages et la dynamique des interactions vers une forme collaborative, en observant les relations de réplique dans les fils de discussion. Les réseaux de communication, principalement orientés vers la dissémination de conseils, peuvent évoluer vers des relations de discussion lorsque des interactions collaboratives entre conseillers se développent. Les formes de la distribution évoluent si les couplages entre messages et les interactions entre auteurs évoluent en même temps. Du point de vue de l'apprentissage social, le couplage Q/R change, en effet, dès que les conseillers discutent entre eux. Dans une véritable logique de discussion, le rôle du questionneur doit aussi pouvoir évoluer : A doit pouvoir devenir informateur et produire, ~~soit des réponses, soit des répliques évaluatives~~ (cf. Annexe C). Pour que le système évolue, il faut donc que les formes de communication et le mode d'interaction entre les agents puissent coévoluer.

Lorsque le fil s'étend au-delà des trois messages, ces configurations élémentaires s'intègrent le plus souvent sous la forme d'un graphe hybride d'abord en forme d'éventail et ensuite en forme filiforme (exemples 1 et 2, Fig. 7).

<sup>21</sup> Cette triade exprime bien la structure novice/deux experts (N/E2), analysée par A. Goldman, dans son analyse du problème de l'expertise, lorsqu'un novice N reçoit le conseil de deux experts (Goldman, 2002, pp. 139-181). Le problème N/E2, selon A. Goldman, est que le novice se trouve dans la position d'avoir à évaluer deux réponses d'expert à partir de sa propre opinion.

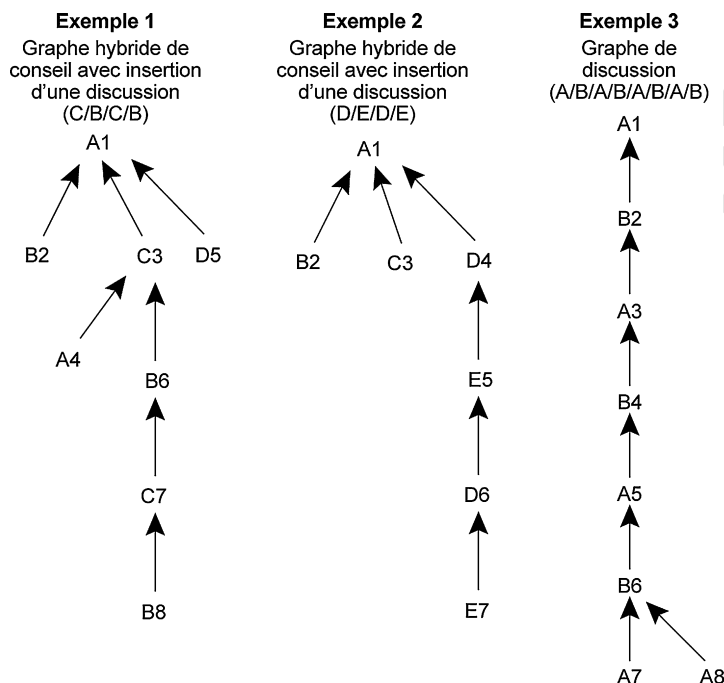


Fig. 7. Exemples de configurations relationnelles au sein d'un fil de discussion de plus de trois messages.

Dans l'exemple 1 de la Fig. 7, le graphe commence, en effet, comme un graphe de conseil avec une forme en éventail de largeur 3 et ensuite devient un graphe de discussion qui prend une forme filiforme avec réintervention de A, de B et C. Le taux d'interaction entre les contributeurs est fort dans ce fil, puisque le rapport entre nombre d'auteurs/nombre de messages est de 4/8 (cf. Annexe A).

Dans l'exemple 2 de la Fig. 7, le graphe prend aussi d'abord une forme en éventail de largeur 3 et devient ensuite un graphe à forme filiforme à partir de la troisième réplique. Le taux d'interaction est plus faible dans ce fil, puisque le rapport entre le nombre d'auteurs et le nombre de messages est de 5/7 (cf. Annexe B).

Maintenant lorsque le fil s'étend, il peut se présenter un cas plus rare où le graphe est un pur réseau de discussion s'il prend une forme purement filiforme et loge une interaction collaborative dialogique entre deux intervenants. Ainsi dans l'exemple 3 de la Fig. 7, le graphe de discussion prend une forme filiforme quasi pure. Il est de largeur 2 à cause de la forme que prend la clôture du fil où A intervient deux fois. Le taux d'interaction est donc très élevé dans ce fil, puisque le rapport entre le nombre d'auteurs et le nombre de messages est de 2/8 (cf. Annexe C).

## 6. Conclusion

La distribution des connaissances au sein de la liste des utilisateurs de *Debian* se présente comme un système massivement distribué d'expertise collective, dont on peut examiner la dynamique en observant l'évolution des réseaux de communication sur des périodes longues. La transformation d'une relation de conseil en relation de discussion passe par une modification de la structure des fils de discussion dont l'organisation favorise au départ le formatage des mes-

sages en questions et en réponses. L'observation des flux d'échanges sur la liste *Debian.user:french* montre que le développement de relations de discussion qui conduit à une atténuation du rôle du questionneur est en tension avec le développement des relations de conseils. Mais si les protocoles *Usenet* donnent la préséance à la relation de conseil, ils n'interdisent pas pour autant l'émergence de relations de discussion. C'est dans cette mesure que le système fonctionne comme un système distribué ouvert où le couplage entre les composants, artificiels et humains, n'est pas défini d'avance.

Les réseaux épistémiques évoluent lorsqu'un conseil donne lieu à une discussion. La possibilité qu'une relation de conseil se transforme en relation de discussion est donc une des conditions d'évolution de ces réseaux, car la qualité de la connaissance repose sur l'apparition de couplages qui favorisent une orientation des répliques vers les réponses. Cette modification des réseaux épistémiques peut s'estimer au niveau de la structure locale des fils de discussion et des relations de coréplique, la densité des interactions entre auteurs des messages et l'orientation des répliques vers les réponses étant les deux manifestations principales de cette modification. Sans l'existence d'un mécanisme correctif de mise en discussion des conseils, il n'y aurait pas d'intervention régulière des experts. Ces correctifs créent une dynamique d'extension des coordinations au-delà des dyades questionneur/répondeur car une discussion suppose une augmentation du nombre de messages et une restriction progressive du nombre d'auteurs. Ce changement dans la distribution sociale de la connaissance se traduit par une intensification des interactions sociales orientées vers la collaboration et la discussion des idées. Le renforcement de la discussion favorise un mode de communication fondé sur l'acceptation par l'informateur de ce que ses assertions soient débattues par ses pairs. Pour N. H. Longino (2002), promouvoir une égalité tempérée entre les informateurs crée des circonstances favorables à l'intensification de la discussion collaborative.

Il faut souligner, cependant, les limites de l'approche que nous avons choisie dans le cadre de cet article. Elle devrait être complétée par une analyse de la cognition distribuée au niveau de la communauté car le changement dans les modes de communication accompagne, dans le cas du logiciel libre et de la communauté des *debianistes*, un changement dans les institutions de distribution de la connaissance. La relation locale interindividuelle entre les informateurs, qui est le point de départ de l'analyse des fils de discussion, ne peut rester l'unique échelle d'appréciation des relations. *Debian* se présente comme une communauté numérique qui accueille en son sein une sous-communauté d'informateurs experts et une sous-communauté de questionneurs novices. La détection de ces sous-communautés ainsi que de leur mode d'interaction serait un précieux complément à l'analyse des graphes des fils de discussion<sup>22</sup>.

L'hypothèse « communautaire » en matière de cognition distribuée suggère que la meilleure façon d'associer diffusion et élaboration est d'utiliser la dynamique des interactions entre les agents, en rassemblant novices et experts au sein d'une structure commune. Cette politique épistémique, propre à la culture *hacker* (Himanen, 2001; Tuomi, 2002, 2005), préconise un changement dans la façon dont les informateurs communiquent entre eux et dans la façon dont les novices interagissent avec les experts. Elle s'appuie sur la dynamique distribuée que permettent les technologies réticulaires. Dans une politique classique d'acquisition de connaissance, il serait naturel de séparer les sites où les novices communiquent avec les experts des sites où les experts communiquent entre eux. Or c'est l'existence des nouvelles ressources cognitives fournies par ces

<sup>22</sup> L'analyse des graphes portant sur de grands corpus peut être tout à fait utile pour construire des détecteurs de communautés. C'est un des objectifs que s'est donnée l'ANR « Autograph » dans son programme de recherche.

technologies réticulaires qui rend possible cette conception fortement distribuée de l'acquisition de la connaissance. La possibilité d'accélérer la diffusion des messages, d'augmenter considérablement le nombre d'informateurs consultables, de disposer en ligne de bibliothèques avec une documentation électronique constamment réactualisée et de construire des débats publics entre experts reconnus facilite la création de communautés épistémiques ouvertes et fondées sur la dynamique des interactions entre des informateurs possédant des niveaux d'expertise variés.

## Références non citées

Giere (2002), Thagard (1997), Kirsh (2004), Conein (2004b), Tomasello (1999).

## Remerciements

Nous tenons à remercier Frédéric Aidouni (LIP 6) pour sa contribution décisive au traitement des données statistiques, ainsi que Nicolas Auray (institut TELECOM) pour les discussions enrichissantes sur la communauté des debianistes et sur le phénomène du logiciel libre.

## Annexe A. Graphe hybride de conseil avec insertion d'une discussion (C/B/C/B)

### Divers pb

#### 1 A

Bonjour à tous, j'ai pas de mal de problèmes aujourd'hui dont je n'ai pas trouvé la réponse dans l'archive de la liste, je les soumetts donc à votre sagacité.

- j'ai installé gnome 1.4 en pointant sur la Sid (mais le pb dont je vais parler existe aussi sur la Woody, j'ai essayé) et gdm refuse obstinément de lancer une session gnome, il indique que ceci est impossible !!! Ai-je oublié un paquet essentiel ?  
présents : tous les paquets téléchargés lorsqu'on fait apt-get install gnome-applets et apt-get install sawfish-gnome ;
  - j'ai récemment essayé Postfix, tout fonctionne bien, fetchmail récupère le courrier de tous les utilisateurs à chaque connexion et postfix le place dans/var/mail/utilisateur. Là où intervient le pb est lorsque je veux utiliser procmail pour ensuite trier ce courrier et le placer dans utilisateur/Mail/boite ce que je faisais facilement avec Exim (et un .forward). Postfix appelle procmail (je l'ai vérifié dans les log) (c'est en fait procmail qui place le courrier dans/var/mail/utilisateur) mais impossible de faire lire par procmail les fichiers/etc/procmailrc ou utilisateur/.procmailrc. J'ai lu pas mal de doc sur la rédaction de ces fichiers mais rien à faire ;
  - je veux faire du DNAT et du SNAT, j'ai bien lu les HOWTO et tout fonctionne bien lorsque je tape les commandes. Pour les obtenir à chaque démarrage dans quel fichier faut-il les placer/etc/rc.local n'existant pas ? Une autre solution est d'utiliser ipmasq (en le modifiant un peu) comme je le faisais avec les noyaux 2.2.\* mais ce paquet ne semble pas fonctionner alors qu'en lisant les scripts je me suis rendu compte que iptables est en principe bien appelé.
- Voilà, cela fait beaucoup, je vous remercie donc pour vos réponses.

#### 2 B

Serait-il possible d'avoir plus de détail ? Il dit juste que c'est impossible ? Pas d'autres messages ?

3 C

Tu peux créer un script/etc/init.d/nat qui fait ton boulot, et créer les liens symboliques qui vont bien dans/etc/rc<runlevel>.d/ (commande update-<jesaisplusquoi>). Ton script doit comprendre un argument qui peut être « start » ou « stop ».

4 A

Gagné, il me manquait le paquet gnome-session, merci pour vos réponses (en particulier Christian).

D 5

gdm n'installe pas gnome-session mais le recommande, ce qui fait que si tu passes par apt-get gnome-session n'est pas installé. En revanche, il est installé si tu utilises dselect.

Moralité, dselect est mieux qu'apt-get.

Ce n'est même pas la peine de faire un rapport de bug.

Christian

6 B

Y'aurait pas une petite option à apt-get pour installer les paquets recommandés ?

7 C

Si elle existe je ne la connais pas et j'ai rien trouvé dans la page de manuel.

Christian

8 B

Procmail refuse de fonctionner si les permissions/propriétaires du.procmailrc sont louches. À vérifier, chez moi, c'est -rw-r--r-- 1 cgo cgo 5844 août 23 09:23 .procmailrc pour le compte cgo.

Idem pour le .forward.

Charles

Annexe B. Graphe hybride de conseil avec insertion d'une discussion (D/E/D/E)

aie ! upgrade difficile

1 A Rien ne va plus pour les applis GTK. -(((Sous kde ou les applis tk, tout passe, mais gimp ou le gnomecc remplace tous les caractères affichés par des petits rectangles. Quelque chose

2 B Bah, chez moi il a suffi d'un redémarrage d'X pour que tout se remette en ordre...

3 C Apres, redémarrage de X et tout va bien...

4 D Vérifier que le charset est bien mis dans les sélections de fontes, il y en a de nouveaux. De même je ne suis pas sûr que tous les packages de fontes fassent un xset fp rehash après l'installation.

5 E Aucun paquet ne fait un xset (faire un grep xset/var/lib/dpkg/info/xfonts\*) C'est un bug ou c'est volontaire ?

6 D Aucun paquet ne fait un xset (faire un grep xset/var/lib/dpkg/info/xfonts\*) C'est un bug ou c'est volontaire ? Je dirais bug, il faudrait au moins le faire quand \$DISPLAY est un display local (voire quand \$DISPLAY est mis, faire un fp rehash ne peut causer aucun dégât).

7 E Si je me souviens bien de mes expériences avec Sawfish, il est impossible d'obtenir la variable DISPLAY dans les scripts d'installation (elle est toujours à : 0.0) Quelqu'un a vu des rapports de bug à ce sujet ?

Annexe C. Graphe de discussion (A/B/A/B/A/B/A/B)

Star office 5.2/Laser Jet 4L



- 1 A Comment faire fonctionner cette imprimante sous SO 5.2 (Debian 2.2)  
Dans le Printing-HOWTO, ils précisent que toute imprimante fonctionnant bien sous Linux fonctionne(rail) bien avec SO...  
La 4L fonctionne parfaitement ailleurs que sous SO
- 2 B Tu choisis "generic printer" (c'est à la lettre G, ce qui est assez idiot AMA) et il va l'envoyer en PS à l'imprimante. Comme la 4L est PS, pas de Pb. Et puis si elle ne l'était pas, tu aurais installé des filtres pour imprimer du postscript, donc pas de Pb non plus. En fait, Soffice sous Linux ne gère pas du tout l'impression
- 3 A La 4L n'est pas PostScript  
« si elle ne l'était pas, tu aurais installé des filtres pour imprimer du postscript, donc pas de Pb non plus »  
Non. Cela marche avec magifilter et ghostscript pour imprimer directement à la main mais le passage par SO est catastrophique. Manifestement, il court-circuite l'impression du système. En fait, j'ai essayé l'impression dans un fichier puis à l'imprimante. Le fichier est bien PostScript mais si j'essaie de le visualiser avec gv, j'ai le message d'erreur suivant :  
error : invalidfont in findfont  
base fontdict : time roman font...  
Késako ? Manque-t-il un paquet sur le système ?
- 4 B Ouups. Je m'a trompé. C'est pas des 4L qu'on a ici. (c'est pas des 2CV non plus (je sais, elle est nulle)). Mais ce sont des postscripts et, effectivement, ça marche. Je pense plutôt que c'est SO qui merdoit la génération de PS, avant de passer par lpr (vérifie dans la config de l'imprimante que la commande d'impression est bien lpr -P...). Mais je peux me tromper. Quand j'imprime dans un fichier ps (en utilisant l'imprimante « Generic »), je peux visualiser le ps sans pb avec gv (preuve : <http://www.limsi.fr/Individu/nico/debian/download/test.ps>) Si tu veux, envoie-moi le ps que tu génère. Si je n'arrive pas à le visualiser, c'est que ton SO est mal installé.
- 5 A C'est marrant parce qu'ici, au boulot (Solaris), je ne peux lire ton truc avec ghostview: Warning: failed to allocate 5 × 5 × 5 RGB cube. Error: /undefined in setpagedevice  
Operand stack: -dicttype- Execution stack: %interp.exit -nostringval- -nostringval- -nostringval- false -nostringval- -nostringval- -nostringval-Dictionary stack: 530/547 0/20 6/200 60/200  
Ghostscript: Unrecoverable error, exit code 1 Et tu as une version 3.0 du ps comme pour SO.  
Quelle est ton interpréteur de ps ? Avec quoi imprimes-tu ? Quelle version de Debian ?
- 6 B Normal, c'est du SO. Je suis en « stable » (2.2 mise à jour en ligne). Mon interprète PS es "gs 5.10-10.1". J'imprime sur différentes imprimantes (plusieurs files) toutes PS  
J'espère que ça t'aide,
- 7 A Après vérification, le PostScript qui passe bien avec ghostview est le 2.0 (essai avec du TeX ≥ dvi ≥ ps). Dans l'en-tête du fichier de SO, il s'agit de PostScript 3.0. Si cela eût mettre quelqu'un sur la piste...
- 8 A Salut,  
Bon, j'ai la solution. Désolé, mais je traitais le cas par téléphone et je n'avais pas la bécane sous la main. Dans ce cas précis, le pilote d'imprimante générique suffisait si l'impression est correctement configurée. Manque de bol, pour gs, il manquait les paquets gsfonts et gsfont-x11, ce qui faisait merder l'impression de SO et pas celle en ligne de commande Cela se serait manifester tôt ou tard en ligne de commande (éternel problème de la gestion des recommandés avec apt). Toutefois, Bernard m'a refilé (entre autre) le tuyau suivant – Tout sur StarOffice en français <http://www.staroffice.online.fr> On y trouve dedans la FAQ en français. Bien plus complet que le site de SUN. On y parle justement du problème de PostScript 3.0. La version courante est boguée et il y a un patch à télécharger (attention, 14 Mo...). Dans mon cas, ce n'était pas nécessaire. Il préconise quant à eux l'emploi de cups pour l'impression  
Merci à Hervé P., Nicolas S., Bernard S. et Christian M. pour leur aide

## Références

- Auray, N., 2005. Le sens du juste dans un noyau d'experts : *Debian* et le puritanisme civique, In: Conein, B., Massit-Folléa, F., Proulx, S. (Eds.), *Internet : une utopie limitée*. Presses de l'université de Laval, Sainte-Foy (Québec), pp. 71–94.
- Clark, A., 1997. *Being there: putting brain, body and world together again*. MIT Press, Cambridge, Ma.
- Clark, A., Chalmers, D., 1998. *Extended mind*. *Analysis* 58, 10–23.
- Conein, B., 2004a. Communauté épistémique et réseaux cognitifs : coopération et cognition distribuée. *Revue d'économie politique*, *Marché en ligne et communautés d'agents*, 141–60.

- Conein, B., 2004b. Relations de conseil et expertise collective : comment les experts choisissent leurs destinataires dans les listes de discussion. *Recherches sociologiques, connaissance et relations sociales*, 61–74.
- Conein, B., 2004c. Cognition distribuée, groupe social et technologie cognitive. *Réseaux* 124, 55–79.
- Conein, B., 2006. Technologie de la communication et épistémologie sociale : comment les TIC favorisent l'acquisition de connaissance. In: Bouvier, A., Conein, B. (Eds.), *L'épistémologie sociale : une théorie sociale de la connaissance*. Éditions de l'EHESS (raisons pratiques), Paris.
- Davis, S., 2002. Conversation, epistemology and norms. *Mind and Language* 17 (5), 512–537.
- Dennett, D., 2000. Making tools for thinking. In: Sperber, D. (Ed.), *Metarepresentations: a multidisciplinary perspective*. Oxford University Press, New York.
- Dorat, R., 2005. Modélisation des threads de discussion dans une liste de diffusion. Mémoire de DEA, université Denis-Diderot, Paris-7, juin 2005.
- Foray, D., Zimmerman, J.-B., 2001. L'économie du logiciel libre, organisation coopérative et incitation à l'innovation. *Revue économique, économie de l'Internet* 52, 77–93.
- Gensollen, M., 2004. Biens informationnels et communautés médiatées. *Revue d'économie politique, marché en ligne et communautés d'agents*, 9–40.
- ~~Giere, R., 2002. Distributed cognition in epistemic cultures. *Philosophy of Science*, 637–644.~~
- Giere, R., Moffatt, B., 2003. Distributed cognition: where the cognitive and the social merge. *Social Studies of Science* 33 (2), 1–10.
- Goldman, A., 1999. Knowledge in a social world. Clarendon Press, Oxford, 161–188.
- Goffman, E., 1961. Encounters, two studies in the sociology of interaction. Mac Millan, New York.
- Himanen, P., 2001. L'éthique hacker et l'esprit de l'ère de l'information. Exiles, Paris.
- Hollan, J., Hutchins, E., Kirsh, D., 2000. Distributed cognition: toward a new foundation for human–computer interaction research. *ACM Transactions on Computer-Human Interaction* 7 (2), 174–196.
- Hutchins, E., 1995a. Cognition the wild. MIT Press, Cambridge, MA.
- Hutchins, E., 1995b. How a cockpit remembers its speeds. *Cognitive Science* 19, 265–288, traduction française, *Sociologie du travail, Travail et cognition*, 36 (4), 451–473.
- Kirsh, D., 1999. Distributed cognition, coordination and environment design. *Proceedings of the European conference on Cognitive Science*, pp. 1–11.
- Kirsh, D., 2004. Metacognition, distributed cognition and visual design. In: Gardinfas, P., Johansson, P., Erlbaum, L. (Eds.), *Cognition, education and communication technology*, pp. 147–180.
- Kirsh, D., 2006. Distributed cognition: a methodological note. *Pragmatics and Cognition* 14 (2), 249–262.
- Lakhani, K., von Hippel, E., 2000. How Open Source software works. MIT Sloan School of Management, Working paper no. 4117.
- Latapy, M., Dorat, R., Conein, B., Auray, N., 2007. Multi-level analysis of an interaction network between individuals in a mailing list. *Annales des télécommunications, analyse des trafics et traces d'usages*, 325–349.
- Lazega, E., 2001. The Collegial phenomenon. Oxford University Press, Oxford.
- Longino, N.H., 2002. The Fate of knowledge. Princeton University Press, Princeton.
- Marcoccia, M., 2004. L'analyse conversationnelle des forums de discussion : questionnements méthodologiques. *Les carnets du CEDISCOR* 8, 23–38.
- Norman, D., 1991. Cognitive artifacts. In: Carroll J. M. (Ed.), *Designing interaction: psychology at the human–computer interface*. Cambridge University Press, New York, traduction française, 1993, Éditions de l'EHESS (raisons pratiques), Paris.
- Sperber, D., Claidière, N., 2006. Why modeling cultural evolution is still such a challenge ? *Biological Theory* 1 (1), 20–22.
- ~~Thagard, P., 1997. Collaborative knowledge. *Nous* 31 (2), 242–261, traduction française, 2006, Éditions de l'EHESS (raisons pratiques), Paris.~~
- Tomasello, M., 1999. The cultural origins of human cognition. Harvard University Press, Cambridge.
- Tomasello, M., Ratner, H., 1993. Cultural learning. *Behavioral and Brain Sciences* 16, 495–552.
- Tuomi, I., 2005a. Internet, innovation and Open Source. *First monday, Peer-reviewed Journal on Internet*.
- Tuomi, I., 2005b. The future of Open Source. In: Wynants, M., Cornelis, M. (Eds.), *How open is the future*. VUB University Press, Brussel, pp. 429–459.
- Q4 Welser, H., Gleave, E., Fisher, D., Smith, M., Visualizing the signatures of social groups in discussion groups. *Journal of Social Structure* 8, à paraître.