

Contribuer n'est pas collaborer : un focus sur les dispositifs d'annotation de documents audiovisuels.

Vincent PUIG

Le Web 2.0 que nous connaissons aujourd'hui est essentiellement un web de conversation, d'échanges et de contributions mais rares sont encore les sites proposant de réels outils de collaboration. Nous voudrions montrer dans cet article que ce sont trois modes d'individuation psychique et collective ¹ qui mobilisent des organologies² différentes. Nous prendrons comme critères d'analyse pour étudier cette organologie, quelques uns des principaux enjeux de recherche de l'IRI : 1) la synchronisation des flux et la combinaison de différents régimes d'attention, 2) la combinaison de métadonnées institutionnelles et amateurs (*top-down* et *bottom-up*) dans des espaces critiques voire polémiques où la visualisation des contributions est un point clé, et 3) la place nécessaire de la sensorimotricité et de l'engagement du corps.

DISTINGUER CONTRIBUTION ET COLLABORATION

A bien des égards, la lecture/écriture de textes est encore le modèle dominant du Web que nous connaissons aujourd'hui, nous verrons plus loin que cependant le Web devient largement audiovisuel en volume (80% en 2010), en nombre de documents, mais surtout en mutation de nos pratiques en temps d'activité et plus

¹ SIMONDON Gilbert, *L'individuation psychique et collective*, Préface de B. Stiegler, Aubier, 2007

² <http://www.arsindustrialis.org/glossary>

profondément en terme de culture partagée. Si l'on analyse la quantité de textes lus, la proportion de lecture humaine « individuelle » est désormais inférieure à la « lecture collective » par les machines et spécialement celle des moteurs de recherche qui filtrent, amplifient, nous revendent *nos* propres lectures, et par là les rendent collectives³. Qu'en est-il de l'écriture ? Nous héritons largement d'une histoire de l'écriture en tant qu'acte individuel, symbolique, intime mais aussi au fondement de la majorité de la production intellectuelle contemporaine par le truchement de la presse et de l'édition, y compris sur le Web. Pourtant, on ne peut les séparer des écritures collectives qui ont toujours existé, des travaux des copistes à la rédaction des textes législatifs, et qui sont aujourd'hui majoritaires si l'on inclut la production de textes par le monde industriel (contrats, projets, documents de communication, sites Web, ...). Mais les outils d'écriture collaborative reposent encore largement sur la ré-écriture ou l'écriture répartie (chacun prenant en charge une partie du document), et beaucoup plus récemment sur l'écriture séquentielle, chacun pouvant à la fois ré-écrire ou ajouter comme dans les plus récents traitements de texte avec suivi de modification. On aurait pu penser que l'avènement du Web et notamment du Web collaboratif (ou 2.0) allait largement bouleverser ce paysage de l'écriture collective. *Not yet*. En effet, les grands sites « collaboratifs », qu'ils soient centrés sur l'écrit comme *Wikipedia* ou sur l'image (animée ou non) tel *Flickr*, *YouTube* ou *Dailymotion*, sont en fait largement contributifs mais encore très peu collaboratifs et ceci pour des raisons sociales et culturelles⁴ mais aussi dans une large mesure pour des raisons technologiques. Les outils proposés sur ces plateformes sont en effet de plus en plus polarisés sur la facilité à charger un contenu, à le rendre visible, à le partager, voire le commenter. Ces outils encouragent le développement d'activités soit disant collectives qui camouflent en fait une grande solitude et de

³ GIFFARD (Alain), STIEGLER (Bernard) *Pour en finir avec la décroissance*, Flammarion, 2009

⁴ cf. la proportion des 90-9-1 un peu trop rapidement érigée en règle

l'individualisme. Les outils réellement collaboratifs, au sens où ils devraient permettre d'écrire ensemble ou plus largement de produire un espace réellement symboliquement partagé, sont encore rares sur le Web et encore plus rares dans le champ de l'écriture de documents audiovisuels. A l'IRI, nous désignons ces outils comme des outils d'individuation collective⁵ ou de transindividuation, parfois en opposition avec les outils d'individualisation, de personnalisation ou de *profiling*. Sur un site comme *Wikipedia*, qui possède probablement les outils les plus élaborés dans ce domaine, se mêlent étroitement et de façon très profonde les fonctions de contribution (rédiger un article) et les fonctions de collaboration qui articulent d'ailleurs encore assez mal l'espace de discussion et la notice. La pertinence de l'outil va de fait dépendre : 1) de l'unité de sens (le niveau auquel une modification ou une discussion peut prendre place), 2) de l'interface de discussion et de modification qui pour le texte s'apparente encore largement à de l'écriture dans la marge d'un livre⁶, 3) du dispositif de visualisation des contributions, de leur évolution dans le temps, dans l'espace et dans toutes les dimensions sémantiques imaginables (projets et ateliers dans *Wikipedia*). Cependant, même un géant comme Google qui a partiellement tenté de repenser ce problème avec *Google Wave* n'a pas encore réussi à convaincre alors même que l'outil combine discussion et édition sur un même espace synchrone.

⁵ SIMONDON (Gilbert), *L'individuation psychique et collective*, Préface de B. Stiegler, Aubier, 2007

⁶ voir logiciel Co-ment de SopinSpace

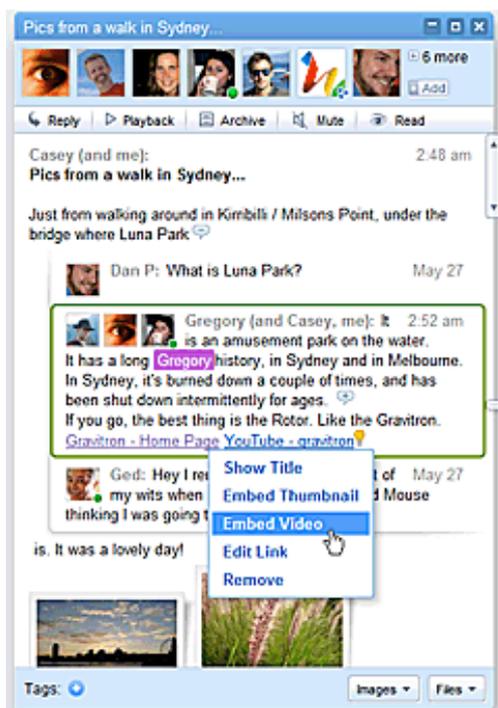


Fig. 1 : Insertion d'une vidéo dans un texte collectivement édité en temps réel dans GoogleWave

DYNAMIQUE DE L'AMATEUR ET COMBINAISON TAXONOMIES/FOLKSONOMIES

Le premier critère déterminant pour distinguer le développement de réels espaces collaboratifs réside selon nous dans la capacité à mobiliser une communauté dans un espace de confiance dont l'outillage – règles, interfaces, vocabulaire et métadonnées sous-jacentes – est non seulement maîtrisé par les utilisateurs mais constitue en lui-même un critère d'identification. Ceci est un point particulièrement important dans l'histoire des communautés

d'amateurs, des eaux-fortes du Comte de Caylus⁷ à la caméra Super 8. L'amateur 2.0 souhaite aujourd'hui s'approprier et configurer des interfaces numériques dont il partage la maîtrise et la connaissance intime avec d'autres amateurs. Un exemple original de cette dynamique est développé par le Studio Brocéliande avec son projet de plateforme de dessin animé participatif qui mettra à la disposition des amateurs des contributions professionnelles à librement reprendre, modifier, interpréter⁸.



Fig. 2 : Interface du prototype THD de dessin animé participatif

Dans un autre genre, celui du jeu vidéo, l'on sait que les outils de création professionnels sont de plus en plus souvent mis à disposition des joueurs pour finaliser les créations des studios (en supprimant les *bugs* restants), mais aussi pour améliorer ou augmenter certains aspects du jeu et, du même coup, en allonger la durée de vie. La pollinisation s'opère sur les différents forums consacrés aux échanges, du simple partage d'information à la

⁷ Voir le séminaire IRI "La figure de l'amateur" sous la direction de Jacqueline Lichtenstein, <http://www.iri.centrepompidou.fr/evenement/museologie>

⁸ <http://www.portailthd.fr/tribes/tribe/MESSANN/>

création en équipes de *patches*⁹ ou de *mods*¹⁰ en tous genres. Quitte à devenir l'un des principaux attraits du jeu, comme l'ont bien compris les créateurs de *Little Big Planet* qui propose aux joueurs de bâtir eux-mêmes des niveaux à partir d'un éditeur. Pensons aussi au phénomène *Minecraft*, jeu indépendant vendu (fait exceptionnel pour une création de ce type) à près de deux millions d'exemplaires en l'espace de quelques mois, permettant de constituer son univers et de le partager avec d'autres en ligne. Jason Rohrer, lui-même créateur reconnu dans ce domaine, s'est emparé du concept, comme n'importe quel joueur peut le faire, afin de créer son propre univers, *Chain World*. Celui-ci tient sur une clef USB qui passe de mains en mains, le partage de l'univers numérique étant conditionné par la rencontre en face à face.

Ainsi donc, dans le contexte de l'économie de la contribution, les amateurs constituent une dynamique qui « pollinise » tous les domaines notamment par la production massive de métadonnées non contrôlées (folksonomies). Pour que cette dynamique conduise à la construction de vrais espaces collaboratifs, des méthodes de croisement avec les systèmes structurés (taxonomies) est selon nous un critère déterminant. C'est tout l'enjeu du projet FUI CineCast¹¹, initié par l'IRI, qui réunit la BNF, la Cinémathèque, le Forum des images, la BPI et l'Ina sur la question des réseaux sociaux cinéphiles et des outils permettant de tisser des liens avec les fonds d'archives cinématographiques. Pour parvenir à cette convergence l'IRI propose notamment la fabrication dynamique d'ontologies à l'aide de modèles pour le tagging tels que *NiceTag*, dans une optique la moins contraignante possible pour le contributeur, afin de lui permettre

⁹ Les patches sont des portions de code destinés à apporter des modifications à un logiciel et, par extension, à un jeu.

¹⁰ Un mod est une modification apportée à un jeu ou un jeu autonome créée à partir d'un autre qui lui sert de support.

¹¹ <http://www.cinecast.fr>

d'exprimer l'intentionnalité de son acte de tagging ou la relation entre son tag et une ou plusieurs ressources¹².

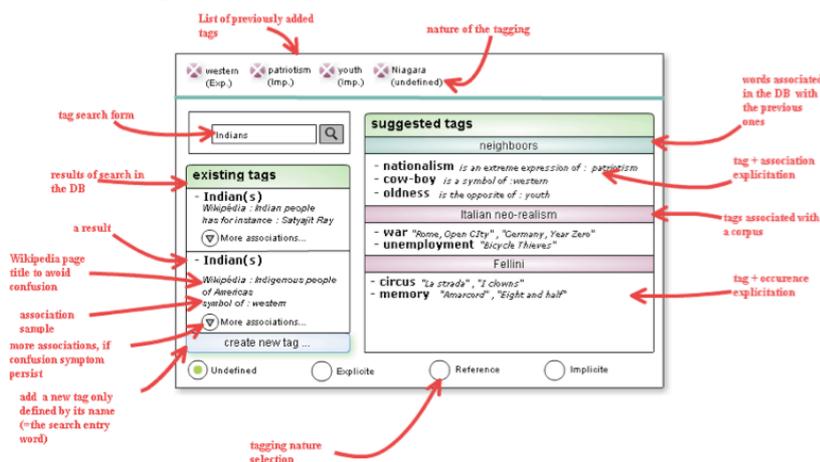


Fig. 3 : maquette d'interface de tagging de films alimentant des graphes de concepts (Damien Pignaud, IRI-Un. de Tokyo-Hitachi systems, 2011)

FLUX DE CONTRIBUTION SYNCHRONES OU ASYNCHRONES ET DIFFÉRENTS RÉGIMES DE L'ATTENTION

Le second critère différenciant les outils de collaboration porte selon nous sur la capacité à synchroniser les flux de contribution. Comme on l'a rappelé avec *GoogleWave*, l'approche synchrone privilégie la conversation à la manière d'un « chat » et donc un régime d'attention dispersé (ou hyper-attention pour reprendre la terminologie de Katherine Hayles) aux dépens d'un régime d'attention soutenue (ou *deep-attention*) celui-là même qu'on trouve en situation d'écriture ou dans *Google Docs* (document éditable à

¹² LIMPENS, (Monnin), LANIADO (Gandon), « Speech acts meets tagging »: NiceTag ontology. Dans *Proceedings of the 6th International Conference on Semantic Systems and the 5th International Conference on Pragmatic Web*. 2010

plusieurs en quasi-synchrone). Dans le champ audiovisuel, cette question de la synchronisation des flux de contribution est plus récente mais particulièrement critique. Il s'agit en effet de synchroniser deux flux temporels, celui du média diffusé et celui des contributions. Pour cela on donnera ici trois illustrations différentes de l'équilibre qui peut être trouvé entre synchrone et asynchrone. Poussé à l'extrême vers la synchronisation la plus directe, on mentionnera le site vidéo japonais *Nico Nico Douga* qui permet d'afficher son tag directement sur la vidéo. Très populaire au Japon, ce système constitue une sorte d'instrument de tagging en temps réel qui procède à la fois de la dynamique du tagueur de rue et de la dynamique du jeu collectif. De fait, le défilé des tags (chacun sa ligne) finit par prendre plus d'importance que la vidéo elle-même et constitue par conséquent une création collective temps réel (fig. 4). A son tour, cette production effrénée de métadonnées encourage dans une très large mesure la création de contenus originaux qui susciteront des réactions en masse (le partage d'une vidéo étant l'occasion d'ouvrir un nouvel espace d'échange tendant à s'autonomiser : la donnée originelle s'efface derrière la métadonnée dont elle devient le support voire le prétexte).



Fig 4 : Commentaires en temps réel sur une vidéo dans le système Nico Nico Douga

Dans un mode où les commentaires ne sont pas produits simultanément mais qui privilégie la contribution orale sur une vidéo, il faut citer aussi le système *VoiceThread* qui est né dans le

contexte des vidéo amateurs et notamment pour commenter les films de famille (fig. 5).



Fig. 5 : contributions vocales sur un film de famille (VoiceThread)

Dans le champ des médias, et notamment dans le contexte de la TV connectée ou des dispositifs multi-écrans (TV + mobile ou tablette), plusieurs systèmes de synchronisation sont proposés¹³ mais *Twitter*, largement répandu comme outil de commentaire en direct dans le contexte des conférences, s'impose dès à présent pour commenter le flux TV comme l'a récemment illustré le Monde.fr pour l'intervention de Nicolas Sarkozy du 10 février 2011 mais aussi pour alimenter de multiples dispositifs de visualisation d'activité¹⁴. En novembre 2009, l'IRI a synchronisé les 1500 tweets produits pendant les *Entretiens du Nouveau Monde Industriel* à l'enregistrement de la conférence. L'objectif était de s'appuyer sur une dynamique de contribution synchrone pour indexer un document vidéo navigable et annotable dans l'environnement Lignes de temps (fig. 6) et *in fine* permettre au contributeur de participer à la construction d'un espace critique voire polémique.

13 <http://devantlatele.com>, *Twitt"TV*

14 <http://www.interfacesriches.fr/2011/02/03/tour-dhorizon-des-outils-de-visualisation-de-tweets/>

synchrone ce qui, les expériences l'ont prouvé, stimule et augmente encore le nombre de contributions. Le dispositif est actuellement en test pour France Culture, où il intègre le partage sur les réseaux sociaux, la visibilité par les moteurs de recherche au niveau du segment (norme W3C *Media Fragments URI*) et la possibilité d'annotation *a posteriori*. De nouveaux modes de représentation du débat sont également testés notamment avec l'outil *DebateGraph* (Fig. 8) qui permet d'articuler les points de débat pour naviguer à la fois dans la polémique et dans son expression dans l'enregistrement.

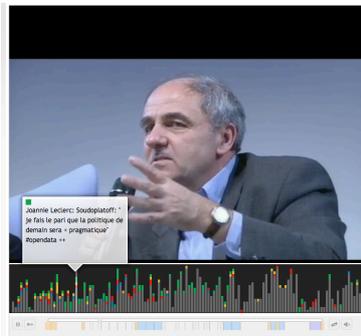


Fig. 7 : Navigation dans la vidéo par les couleurs polémiques du débat tout en affichant les tweets (ici un tweet en adhésion)

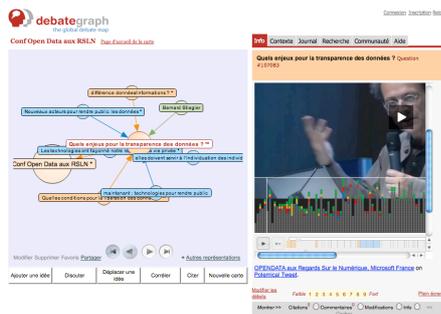


Fig. 8 : Manipulation du graphe reliant les prises de position et leurs relations polémiques

PLACE DE LA SENSORI-MOTRICITE DANS LES FUTURS ESPACES COLLABORATIFS

Pour conclure nous aimerions évoquer le ressort déterminant de la contribution relatif à l'implication du corps. A la suite des travaux de Francisco Varela sur l'énaction, de nombreuses recherches en sciences cognitives¹⁶ ont démontré l'importance des boucles sensori-motrices dans la perception des contenus. L'appréhension d'une œuvre, en particulier, passe par sa copie ou, à tout le moins, par la compréhension des outils qui ont permis de la fabriquer. A plus forte raison, en situation de contribution, est-il intéressant de donner accès à des outils de reproduction, fussent-ils des métaphores numériques pour peindre, afin de commenter un tableau¹⁷, chanter pour s'exprimer sur une musique, reproduire un geste de danse pour mieux comprendre une chorégraphie... C'est précisément l'enjeu du projet *Finger's dance* que l'IRI a initié avec Thierry de Mey en s'inspirant de la grammaire de gestes qu'il avait utilisé pour *Musique de table*, puis pour décrire les gestes des danseurs notamment dans le film qu'il réalisait sur la chorégraphie *One flat thing, reproduced* de William Forsythe.

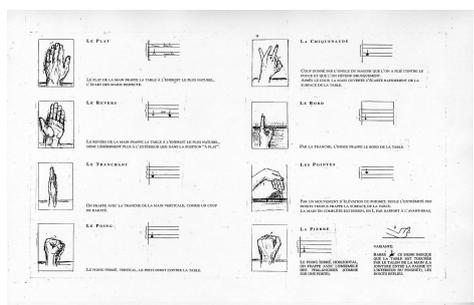


Fig. 9 : partition de gestes pour *Musique de table*, 1987

¹⁶ STEWART (J), GAPENNE (O) *Enaction, Toward a new paradigm for cognitive science*, MIT Press, 2010

¹⁷ le projet <http://www.googleartproject.com/> permet précisément grâce à la haute définition, d'appréhender la dimension tactile, les textures et donc les outils du tableau

Des « transcriptions » des gestes des danseurs ont été conçues et adaptées pour fonctionner sur une tablette tactile de grand format¹⁸. L'expérimentation a montré d'une part la rapidité et l'intérêt de tagger un film de danse avec son geste mais par ailleurs toute la dynamique induite par le processus précisément ici collaboratif puisque les quatre tagueurs pouvaient se répartir les différents moments ou dimensions du film et s'échanger d'un geste leurs tags respectifs. Si cette recherche est née dans un contexte culturel, l'actualité plus récente notamment sur la TV connectée combinée à d'autres écrans (mobiles, tablettes) nous confirme l'enjeu industriel considérable de la normalisation des interactions gestuelles dans des contextes où, malheureusement, la cohérence entre geste de contrôle et geste suivi (ou action déclenchée) n'est pas évidente et, dans la perspective d'une hégémonie industrielle croissante, peut ne pas favoriser le développement de communautés de pratiques et de cercles d'amateurs.



Fig. 10 : tagging gestuel collaboratif de films de danse, déc 2009)