

Journée d'études des URFIST

31 janvier 2007, Paris

« Evaluation et validation de l'information sur internet »

La communauté comme auteur et éditeur : l'exemple de Wikipédia

Laure Endrizzi

chargée d'études et de recherche,
cellule Veille scientifique et technologique, INRP, Lyon

Résumé

L'ensemble des technologies dites 2.0 place l'utilisateur au cœur de la création des contenus numériques tout en l'inscrivant dans une dynamique collective. Ces transformations remettent en cause le modèle éditorial traditionnel, sans offrir de représentations claires et stabilisées des modes de production et de validation qui sont à l'œuvre.

Avec l'exemple de Wikipédia, nous tenterons de comprendre les mécanismes de la régulation éditoriale, pour ensuite nous interroger sur les formes d'expertise sollicitées et les figures de l'auteur.

Nous verrons que les outils et dispositifs qui sous-tendent l'activité s'inscrivent dans un mouvement de maturité croissante et que l'environnement collaboratif favorise l'émulation des énergies créatrices personnelles en misant sur une action critique participante et sur une réflexivité explicite.

Les fonctions qui apparaissent sont significatives d'une tension croissante vers la « professionnalisation », associant diversification et spécialisation des contributions et laissant émerger des rôles somme toute relativement traditionnels.

Ces différents aspects ébauchent les contours d'un véritable modèle éditorial, faisant de Wikipédia, certes un projet collaboratif mais également une encyclopédie à part entière. Des résultats encore fragiles que de futures recherches permettront de confirmer ou d'infirmer.

Avec la faiblesse de ses barrières éditoriales et son actualisation permanente des connaissances, Wikipédia – l'encyclopédie libre que chacun peut améliorer – renouvelle le genre encyclopédique par son approche éminemment démocratique et populaire.

Nombreux sont ceux qui pensent que l'absence d'appel à des auteurs-experts et de validation par les pairs est rédhibitoire à la fois en termes de fiabilité et de crédibilité. Alors qu'affaires de vandalisme et controverses sur les dérives sectaires, marchandes et politiques mobilisent parfois l'attention, Wikipédia s'est installée en deux ans au rang des 12 sites les plus visités au monde selon le baromètre Alexa¹. Et de plus en plus nombreux sont également ceux qui contribuent à la rédaction des articles, dont la croissance exponentielle n'est plus à démontrer. Ces usages, en quelque sorte, confèrent à l'encyclopédie une forme de légitimité, et ce phénomène ne peut déceimment être ignoré, ne serait-ce que parce que les étudiants s'en servent massivement sans être pour autant des utilisateurs « avertis ».

L'avènement du web 2.0, concept popularisé par Tim O'Reilly fin 2005², nous conduit inmanquablement à replacer ces questionnements sur Wikipédia dans un contexte plus large. Ce qui est à l'œuvre ici, c'est le foisonnement des technologies qui placent désormais l'utilisateur au centre de la création de contenus sur internet : ce que les anglo-saxons ont baptisé UGC (*user generated content*)³ – phénomène qui affecte toutes les strates de la production d'information, depuis l'information grand public et journalistique jusqu'à l'information de vulgarisation et la publication scientifique. Ces transformations, génératrices d'incertitudes, modifient en profondeur les modes d'évaluation et de régulation traditionnellement légitimes, sans offrir pour autant de représentations claires et stabilisées sur les processus de production et de validation de l'information.

Avec Wikipédia, la tentation de pointer les articles de qualité médiocre est grande... et facile. Appréhender la dynamique éditoriale dans toute sa complexité nécessite de s'inscrire plus globalement, dans une perspective temporelle. La question n'est peut être pas tant d'ailleurs de savoir si Wikipédia offre des contenus de qualité que de s'interroger sur ce qui constitue une assurance qualité dans un tel processus de production collaborative, comme le rappelle Paul Duguid (2006). Les lois de qualité opérationnelles dans les communautés de logiciels libres peuvent-elles nous aider à comprendre les mécanismes qui sous-tendent l'activité éditoriale ? Celle de Linus, illustrée par la célèbre phrase d'Eric Raymond « *given enough eyeballs, all bugs are shallow* », qui postule qu'au delà d'un certain nombre de contributeurs, le produit gagne en qualité ? Ou bien encore celle de Graham, qui montre que ce qui est de bonne qualité perdure et se développe, alors que ce qui est d'une qualité médiocre tend à être ignoré ou à disparaître ?

Le nombre et la durée jouent-ils véritablement en faveur de la qualité dans le cas de Wikipédia ? Si tous les indicateurs semblent le confirmer, les processus qui sous-tendent l'activité doivent être explorés plus avant. Comment s'opèrent la co-rédaction, la co-validation ? Sont-elles véritablement décentralisées ? Quelles sont les formes d'expertise sollicitées ? Pour quels résultats ? L'examen des dernières recherches depuis notre précédent panorama (Endrizzi, 2006) nous permettra, dans cette contribution, d'apporter quelques éléments de réponse sur le fonctionnement de cette communauté d'utilisateurs qui est à la fois auteur et validateur des contenus produits.

Outils et mécanismes de la régulation éditoriale

Des mécanismes gestionnaires incorporés

Le concept « wiki », développé par Ward Cunningham en 1995, allie immédiateté et simplicité pour permettre à toute personne disposant d'une connexion internet d'en modifier les pages à volonté.

¹ <http://www.alexacom.com/data/details/main?q=&url=http://www.wikipedia.org>

² <http://www.internetactu.net/?p=6421>

³ http://en.wikipedia.org/wiki/User-generated_content

Malgré des fonctionnalités plus ou moins distinctives, tous les wikis offrent différents espaces qui favorisent l'auto-gestion et la régulation a posteriori des contenus : articles, méta-informations et outils de communication sont intimement corrélés au sein d'une même plate-forme.

Ainsi, le MediaWiki, moteur des projets de la fondation Wikimedia, propose divers outils pour :

- suivre l'actualité de la communauté, ses décisions et ses besoins ;
- évaluer son activité interne (statistiques et tableaux de bord) ;
- évaluer son audience externe (revues de presse, bibliographies) ;
- s'informer sur les règles et conventions de participation ;
- participer aux votes et sondages en cours ;
- connaître les rôles et profils des participants ;
- assurer le suivi des articles, normaliser leur structure, etc.

C'est la visibilité et l'interdépendance de ces mécanismes gestionnaires, qui permettent à chacun d'identifier les traces de la co-construction de l'article et du co-développement de la communauté. Sur le plan éditorial, ils participent pleinement à la publicisation des besoins (via les *modifications récentes*, *pages à supprimer*, *projet de la semaine* ou les bandeaux *ébauches*, par exemple) qui elle-même exerce une force d'attraction mobilisatrice pour les contributeurs (« effet piranha »⁴). Sur le plan communautaire, l'incorporation de tels outils favorise l'action critique et la réflexivité explicite, tout en valorisant les interactions sociales (Lawler, 2005) (Wagner, 2006).

Des modes d'appropriation variables et évolutifs

Ces dispositifs communs ne sont cependant pas exploités de manière homogène par l'ensemble des univers linguistico-culturels. Voss (2005) montre par exemple que les articles représentent 60% pour la version allemande et 80% pour la version croate. Autrement dit, les 20 à 40% restants sont composés de fichiers multimédias, pages de discussion, catégories thématiques, modèles pour la standardisation des articles, pages méta dédiées à la gestion et pages personnelles des utilisateurs. Plus la version est jeune, plus la part des contenus encyclopédiques est importante, l'article constituant la cellule de base autour de laquelle d'autres cellules viennent se greffer.

La récente étude conduite sur la version anglaise par les chercheurs de l'IBM Watson research center (Viégas *et al.*, 2007) confirme les variations temporelles dans la croissance de ces différents « namespaces » : entre 2003 et 2005, les articles de l'encyclopédie, multipliés par 9, ont bénéficié d'une croissance nettement moins forte que les pages de discussion des utilisateurs (x 78) et les pages méta dédiées à la gestion du projet (x 68). Le déploiement rapide des *WikiProjects* (collections thématiques d'articles) associé à celui des différents espaces auxiliaires analysés montrent qu'une part croissante de l'activité se recentre sur la coordination plutôt que sur la seule édition d'articles.

Un autre espace a bénéficié récemment d'une croissance forte : les catégories, introduites en mai 2004 dans le MediaWiki, avoisinent fin 2005 les 80000 dans la Wikipédia anglaise, pour couvrir plus de la moitié des articles (Holloway *et al.*, 2006). Le palmarès des 20 catégories les plus utilisées montre une prédominance gestionnaire forte, avec notamment une prolifération de catégories de type « ébauches » (*stubs*)⁵.

⁴ Un article rendu visible, soit pour sa qualité, soit pour sa pauvreté, va concentrer ponctuellement davantage de mises à jour en raison de cette publicisation ; les wikipédiens de la version française nomment ce phénomène l'effet piranha.

⁵ Catégories gestionnaires dominantes : *disambiguation*, *articles_to_be_merged*, *cleanedup_from_october_2005*, *incomplete_lists*, *film_stubs*, *musician_stubs*, *american_people_stubs*, *musical_group_stubs*, *album_stubs*, *politician_stubs*, etc.

Le rôle des pages de discussion

Parmi ces outils, les « *talk pages* » semblent occuper une place privilégiée. Alors que Stvilia et al. (2005) analysent les critères de qualité discutés par les participants dans ces pages, Emigh et Herring (2005) montrent que les échanges qui s'y tiennent sont fortement marqués par le respect de la neutralité de point de vue (NPOV) et participent pleinement à cette tension vers l'« horizon encyclopédique » (Levrel, 2005).

Selon Viégas et al. (2007), les pages de discussion jouent un rôle qui va bien au-delà de la résolution des conflits d'édition mise en évidence dans leurs précédents travaux (2004), contribuant plus globalement à la dynamisation et à la gestion de l'activité d'édition. 50% des messages postés sur ces pages correspondent clairement à une explicitation des mises à jour effectuées ou des améliorations à apporter : elles servent donc à la fois à coordonner et à planifier l'activité. 10% s'apparentent davantage à des demandes d'information de la part de contributeurs qui pointent les faiblesses de tel ou tel article sans véritable intention éditoriale. 8% font référence aux règles et conventions officiellement entérinées par la communauté, assumant alors une fonction pédagogique régulatrice.

Un objectif commun et des principes fondateurs

La dynamique éditoriale s'appuie bien évidemment sur ces outils fournis par le wiki, mais repose également sur quelques principes clés, édictés dès l'origine, qui conditionnent les échanges⁶ :

- « Wikipédia est une encyclopédie » ;
- le respect des règles de savoir vivre dans le processus de rédaction collective (Wikiquette) ;
- le respect du copyright et la licence libre GNU GFDL (general free documentation licence) ;
- la neutralité de point de vue (neutral point of view ou NPOV).

Ce socle commun se décline depuis 2001 dans de multiples règles et recommandations régionales qui, si elles sont significatives de la vitalité organique de la communauté, ne sont pas sans rappeler les conventions pratiquées dans l'édition traditionnelle. Leur explicitation au sein même du wiki ne garantit toutefois pas leur visibilité ni leur appropriation. Alors que cette prolifération de règles n'offre en définitive que peu de prise pour comprendre les mécanismes de l'auto-régulation, plusieurs études s'accordent à montrer que l'articulation entre ces mécanismes gestionnaires flexibles et transparents et ces principes fondateurs joue un rôle structurant suffisant pour réguler l'essentiel de l'activité. Ce n'est que lorsque l'auto-régulation est mise en échec que le processus devient formel, faisant appel à des instances de médiation, voire d'arbitrage (Levrel, 2005), (Reagle, 2005).

Cette maturité progressive des outils et des règles qui guident l'activité éditoriale fait écho au processus d'acculturation mis en évidence par Bryant et al. (2005). Les auteurs montrent comment les Wikipédiens actifs élargissent progressivement leur zone proximale de développement quand certains outils ou espaces deviennent plus visibles parce que plus porteurs de sens. L'activité se ritualise et les rôles se diversifient tout en se spécialisant. Wikipédia, perçue au départ comme une collection d'articles, devient progressivement une collection d'individus entre lesquels des échanges se mettent en place. Ce processus permet l'émergence de formes d'expertise diversifiées. Plusieurs études ont montré que la biographie personnelle joue un rôle clé dans les niveaux et stratégies d'engagement, répondant à une démarche généralement plus pragmatique que contestataire, fondée sur le plaisir de partager et d'apprendre (Levrel, 2005), (Anthony *et al.*, 2005), (Bryant *et al.*, 2005).

⁶ <http://fr.wikipedia.org/wiki/WP:REG>

Formes de l'expertise et professionnalisation

Distribution des pouvoirs techniques et expertise gestionnaire

Les contributeurs sont d'abord caractérisés par leurs statuts au sein de la plate-forme d'édition : *utilisateur anonyme*, *utilisateur enregistré*, *administrateur* ou *développeur*. Les critères d'éligibilité au statut d'administrateur ne sont pas strictement normalisés et peuvent varier selon les versions⁷. Ce sont à la fois l'expérience dont peut justifier le candidat et la confiance qu'il inspire aux membres de la communauté qui sont examinés. De facto, des critères tels que le nombre total d'éditions et leur fréquence au cours des mois précédant la candidature interviennent dans l'instruction. Les votes sont généralement ouverts aux utilisateurs enregistrés et aux administrateurs ; la nomination repose sur le consensus, non sur la majorité : aucun ratio pour / contre n'est fixé a priori.

Le statut d'administrateur s'interpénètre avec des fonctions plus ou moins spécialisées qui dessinent ainsi une hiérarchie de pouvoirs techniques (Levrel, 2005). Seuls les administrateurs sont par exemple autorisés à protéger, supprimer des pages ou bloquer des contributeurs malveillants. Les *vérificateurs d'adresses IP* tracent les contributeurs malveillants. Les *bureaucrates* affectent localement les statuts d'administrateurs et de bureaucrates. Les *stewards* donnent, modifient ou retirent tous les statuts de tous les projets Wikimedia, à l'exception généralement de leur communauté d'origine. L'application de ces pouvoirs reste soumise aux votes des utilisateurs⁸.

Dans la version française, ces fonctions sont réparties de la manière suivante (décembre 2006) : 124 administrateurs, 7 bureaucrates, 7 vérificateurs d'adresses IP, 3 développeurs ; ce qui représente moins de 1% de l'ensemble des contributeurs enregistrés. Et les proportions sont équivalentes dans la Wikipédia anglaise⁹.

Alors que tous les Wikipédiens sont – théoriquement – égaux devant l'activité d'écriture, les tâches qui requièrent une acculturation technique sont donc restreintes à certains utilisateurs. Ils tendent alors à structurer leur contribution sur des tâches répétitives et s'affichent comme porteurs de la tâche pour canaliser les initiatives individuelles. La routinisation des procédures de coopération devient alors plus importante que l'expertise scientifique pour légitimer sa carrière au sein du projet (Levrel, 2005).

Cette routinisation ne s'appuie pas nécessairement sur des pouvoirs techniques spécifiques et prend place dans des groupes aux fonctions plus ou moins formelles. La publicisation de l'activité passe par l'émergence d'un vocabulaire spécifique, signe de la créativité de la culture wiki. Ainsi trouve-t-on dans la Wikipédia française des « *WikiPompier*s » assumant une tâche de médiation dans les conflits, des « *WikiLovers* » chargés de promouvoir les règles de bonne conduite, des « *WikiFourmis* » chargés de l'organisation et de la catégorisation des articles, des « *WikiGnômes* » pratiquant des interventions mineures mais utiles (orthographe, grammaire, désambiguïsation, gestion des liens), des « *WikiFées* » focalisées sur les aspects formels (mise en page, couleur, illustration), des « *WikiGraphistes* », des « *WikiPhotographes* », des « *WikiTraducteurs* », etc.

Expertise éditoriale : normalisation, dynamisation, arbitrage et validation

Si l'expertise éditoriale reste invisible de prime abord, elle peut être décelée localement dans les processus de normalisation, de dynamisation, d'arbitrage et de validation mis en œuvre.

En effet, la normalisation des articles (structure, infobox) et la construction de l'arborescence thématique dans laquelle ils s'insèrent (palettes de navigation, indexation) sont des activités qui

⁷ <http://fr.wikipedia.org/wiki/WP:RfA>

⁸ Par exemple, le vote sur les pages à supprimer dans la version française est ouvert à toutes les personnes disposant d'un compte depuis plus de 3 mois et ayant effectué plus de 150 contributions à l'ouverture du vote.

<http://fr.wikipedia.org/wiki/WP:P%C3%A0S>

⁹ http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Statuts_des_utilisateurs et <http://fr.wikipedia.org/wiki/Special:Listadmins>

requièrent un certain formalisme conceptuel. De même, la mobilisation autour des portails thématiques ou des projets de collaboration fait apparaître une autre forme d'expertise : l'expert ici est celui qui participe à la dynamisation du projet éditorial. Dans une moindre mesure, toute activité visant à rendre publics les besoins d'édition (application des catégories « ébauches » ou de modèles liés au cycle de vie de l'article) contribue à cette dynamisation.

L'expertise éditoriale s'exerce également au niveau de la gestion des conflits, qui, lorsque les stratégies informelles de négociation échouent, nécessite le déploiement d'une médiation, voire d'un arbitrage : des situations extrêmement chronophages qui peuvent en bout de processus conduire à une intervention des administrateurs pour temporairement exclure un utilisateur ou bloquer une page.

La validation pour sa part est basée sur un double processus : la suppression de pages pauvres ou non pertinentes (*PàS*) et la labellisation « articles de qualité » (*AdQ*) ou *Featured articles (FA)*. Les critères justifiant la suppression ou la promotion d'un article s'avèrent diversement élaborés selon les versions, tout en mettant en évidence une maturation progressive de l'activité éditoriale.

Les *PàS* reposent sur un ensemble de règles officielles particulièrement complexes dans le cas de la Wikipédia anglaise, alors que la version française fournit pour l'instant des préconisations essentiellement techniques¹⁰. La validation *AdQ* répond à un processus largement formalisé dans la version allemande, première à avoir publié une version off-line de l'encyclopédie. L'activité des comités de lecture dans les versions anglaise et française n'est pas (encore) systématisée et ne joue qu'un rôle informel dans la labellisation des *AdQ*¹¹ (Riehle, 2006). La définition des standards de qualité dans la Wikipédia anglaise a cependant fait l'objet d'ajustements ; l'application rétrospective des nouveaux critères aux articles *AdQ* a donné lieu à la création d'une catégorie spéciale pour le ré-examen du label : les *Featured articles removal candidates (FARC)* (Stvilia *et al.*, 2005).

Dans la perspective de généraliser les expériences allemandes¹², des réflexions seraient en cours pour introduire des métadonnées sur le statut de l'article et permettre la cohabitation sur le wiki de deux versions : une version « live » telle qu'elle existe actuellement et une version validée comme stable¹³. A ce dispositif technique, pourrait répondre une organisation humaine. La Wikipédia anglaise par exemple s'est dotée d'une « version 1.0 editorial team » chargée de définir et mettre en oeuvre la validation des thèmes et des articles à retenir dans la première version stable (1.0) appelée à voir le jour. Cette équipe canalise notamment les demandes de validation par les pairs sur certaines collections thématiques d'articles¹⁴. Ce projet de cohabitation semble toutefois particulièrement controversé au sein de la communauté et n'a pas donné lieu à des avancées significatives depuis son annonce par Jimmy Wales fin 2005. Selon Florence Devouard, Wikipédia s'orienterait davantage vers une labellisation « article validé » à l'aide de métadonnées facilitant une extraction des contenus pour l'édition de versions off-line, sans pour autant porter atteinte à la dynamique éditoriale¹⁵.

Autorité éditoriale : un nouveau collège invisible ?

Wikipédia se construit au profit d'un contrôle partagé entre tous les participants et remet fondamentalement en cause le modèle de validation par les pairs. Les experts n'ont donc pas de

¹⁰ http://en.wikipedia.org/wiki/Wikipedia:Deletion_policy

http://fr.wikipedia.org/wiki/Aide:Comment_supprimer_une_page

¹¹ <http://de.wikipedia.org/wiki/Wikipedia:Review>

http://en.wikipedia.org/wiki/Wikipedia:Peer_review

http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Comit%C3%A9_de_lecture

¹² commercialisation d'un cdrom et d'un dvdrom, mise à disposition de versions téléchargeables, édition de fascicules thématiques imprimés (« wikireaders »).

¹³ voir l'article « Wikipédia va se dédoubler », paru dans l'édition du 19/12/2005 du journal *Libération*.

¹⁴ http://en.wikipedia.org/wiki/Wikipedia:Pushing_to_1.0

¹⁵ conférence « Wikipédia : serions-nous tous des lumières ? », BM Lyon, 23/01/2007.

<http://www.vive-laculturenumerique.org/index.php/2007/01/29/30-l-essentiel-pour-comprendre-wikipedia>

position a priori privilégiée qu'il s'agisse de la rédaction ou de la validation des contenus. Mieux, la règle « pas de travaux inédits »¹⁶ cadre en quelque sorte leur engagement, recommandant de ne pas utiliser l'encyclopédie pour publier leurs recherches.

NPOV et consensus demeurent cependant des notions éminemment ambivalentes. La moyenne des points de vue est-elle plus pertinente que le point de vue du meilleur des experts ? L'objectivité peut-elle s'établir sur une majorité de contributeurs plus ou moins anonymes, plus ou moins sincères, voire plus ou moins uniques, votant pour telle ou telle vérité scientifiquement non certifiée ? Dans quelle mesure la persévérance – individuelle ou collective (groupe d'influence) – ne prend-elle pas le pas sur la sincérité, conduisant au découragement certains contributeurs ? Comment amateurs et professionnels collaborent-ils à Wikipédia ?

Même si les votes et discussions sont consultables, l'autorité s'exerce donc collectivement selon des modalités peu transparentes et peu systématiques, et reste déconnectée de toute référence à des éléments biographiques extérieurs. Ce fonctionnement méritocratique n'est après tout pas si éloigné de l'édition traditionnelle : ne serait-on pas en train d'assister à l'émergence d'un nouveau « collègue invisible » ? Si les conflits d'édition tendent avec le temps à se diluer dans l'historique des articles (Viégas *et al.*, 2007), quelles sont les formes d'autorité qui conditionnent les arbitrages ? Ces arbitrages sont-ils simplement factuels ou incluent-ils véritablement une dimension éditoriale ? Autrement dit, qui décide de ce qui est important, encyclopédique ? Qui, a contrario, décide de ce qui est anecdotique ? L'emphase sur l'événementiel et l'actualité d'une part, la prépondérance des articles liés aux passions adolescentes d'autre part, ne fragilisent-elles pas la légitimité culturelle de l'encyclopédie ? Dans quelle mesure les aspects gestionnaires ne se substituent-ils pas à l'expertise éditoriale ? La validation ne tend-elle pas à se focaliser à l'excès sur des règles formelles, symptômes d'une bureaucratie rampante ? L'importance d'un critère tel que le nombre de références externes dans la validation *AdQ*, peut laisser songeur quand on sait par ailleurs que le spamdexing (insertion de liens « promotionnels ») est un des principaux fléaux auxquels Wikipédia doit faire face¹⁷... De nombreuses questions que les chercheurs n'ont pas encore investi...

Expertise administrative et politique : cadres et militants

Parallèlement à ces formes d'expertise éditoriale, apparaissent des rôles associés à la gouvernance globale. Certains Wikipédiens cumulent ainsi des fonctions d'administration auprès de plusieurs projets (Wikinews, Wikilivres, Wikitionnaire, etc.) avec des postes plus officiels au sein de la Fondation Wikimedia via le MetaWiki¹⁸ (Riehle, 2006). Leur acculturation « politique » se double généralement de prérogatives techniques : le pouvoir de ces cadres n'est de toute évidence pas négligeable, même s'il reste difficile à évaluer.

Mais les discussions politiques ne sont pas restreintes à cette micro-communauté. Des mouvements partisans plus ou moins structurés émergent çà et là, significatifs de l'engagement de certains Wikipédiens dans le projet éditorial global : les « métapédiens ». Dans la Wikipédia anglaise, ces mouvements prennent la forme de véritables écoles de pensée. A une extrémité du spectre, l'*Association of deletionist Wikipedians (ADW)* rêve d'une encyclopédie expurgée de tout contenu incomplet, malpropre et non pertinent et favorise l'établissement de standards rigoureux. A une autre extrémité, l'*Association of inclusionist Wikipedians (AIW)* s'oppose à une sélection des contenus sur des critères de pertinence, privilégie l'amendement des articles problématiques plutôt que leur suppression et mise sur l'activité organique de la communauté. Entre ces deux extrémités, s'est formée une kyrielle de mouvements intermédiaires, plus ou moins investis d'une mission de réconciliation. Mais ces positionnements ne sont-ils pas mutuellement exclusifs, alors que les uns mettent l'emphase sur l'encyclopédie, les autres sur le wiki (Carr, 2006) ?

¹⁶ http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Travail_in%C3%A9dit

¹⁷ un phénomène de référencement abusif qui devrait être enrayé prochainement, via l'introduction de balises « no follow » pour que les moteurs de recherche ne suivent pas les liens externes.

¹⁸ <http://wikimediafoundation.org/wiki/Home> et http://meta.wikimedia.org/wiki/Main_Page

Wikipédia pourra-t-elle adopter une démarche consensuelle pour résoudre cette querelle historique naissante, ou bien verrons-nous le schisme se matérialiser en deux versions distinctes à moyen terme ? Si ces questions demeuraient marginales dans les premiers temps de la communauté, il est à parier qu'elles occuperont une place grandissante dans les perspectives de standardisation qui sont actuellement étudiées. Le projet Citizendium n'est d'ailleurs pas étranger à ces divergences.

Citizendium = bottom up + top down ?

Citizendium est l'encyclopédie alternative lancée en octobre 2006 par Larry Sanger, co-fondateur de Wikipédia avec Jimmy Wales. Depuis son retrait courant 2002, Sanger (2005) dénonce publiquement le manque de respect pour l'expertise inhérent au projet, considérant l'anti-élitisme comme préjudiciable à la légitimité de l'encyclopédie. Citizendium tire partie de la licence GFDL (general free documentation licence) pour bifurquer : la base intégralement dupliquée continuera à évoluer avec les modifications effectuées sur Wikipédia, alors que la communauté d'experts travaillera sur certains articles. C'est ce qu'on appelle un « fork progressif », la ramification secondaire continuant à se nourrir de la ramification principale, tout en évoluant séparément.

Sanger (2006) justifie son entreprise en défendant l'idée que les experts, en particulier les universitaires, ont un rôle – spécifique – à jouer dans la régulation de l'intelligence collective, même si leurs pratiques ne sont généralement pas caractérisées par cette culture Web 2.0. Une initiative ambitieuse, quasi philosophique, de conciliation entre des logiques bottom-up et top-down, qui n'est pas sans rappeler Nupédia, la grande sœur de Wikipédia, et qui se développe en marge des efforts de standardisation mis en œuvre par la communauté wikipédienne. Un pari qui soulève plus de questions qu'il n'ouvre véritablement de perspectives à ce stade embryonnaire, à la fois sur l'identité des experts, sur la manière dont ils vont exercer cette expertise, sur les contenus qui seront soumis à relecture., etc.

Figures de l'auteur et collaborations

Des auteurs anonymes « bons samaritains »

La proportion des éditions anonymes n'est pas négligeable : 17% dans la version polonaise, 19% dans la version française, 30% dans les versions anglaises et allemandes et 52% dans la version japonaise¹⁹. Si l'anonymat révèle des différences culturelles marquées dans les pratiques, cette proportion n'est cependant guère évolutive, en tous cas dans les corpus considérés : les croissances des utilisateurs enregistrés et anonymes sont en effet parallèles. Si les uns sont (logiquement) plus actifs que les autres, le statut permet-il de repérer des variances dans les stratégies participatives ?

Alors que les relations entre anonymat et vandalisme ne sont pas établies (Viégas *et al.*, 2004), l'anonymat semble pouvoir aller de paire avec une inclinaison plus forte pour les sujets sensibles tels que la drogue et la sexualité (Kolbitsch, 2006a) : ces articles comptabilisent en effet plus de contributions anonymes et plus de contributeurs anonymes uniques par article. Autre variation, moins attendue : la catégorie des pages spéciales (remarquables pour leur notoriété, leur qualité ou leur pauvreté) fait proportionnellement l'objet d'une attention plus soutenue de la part des contributeurs anonymes : un résultat de l'« effet piranha » ?

Plus intéressante est sans doute l'étude d'un groupe de chercheurs du Dartmouth College d'Hanovre (Anthony *et al.*, 2005), qui croisent le niveau de participation et la proportion de contenus retenus après une intervention, pour mettre à jour les styles participatifs des utilisateurs. Les auteurs distinguent les anonymes qui se contentent d'apporter des correctifs mineurs aux articles consultés, de ceux qui produisent très occasionnellement et de manière ciblée des contenus de qualité relatifs à leur domaine d'expertise. Dans cette configuration, le maillon faible est constitué de ces

¹⁹ <http://stats.wikimedia.org/EN/Sitemap.htm> (juillet 2006).

contributeurs anonymes dont la fréquence des interventions est élevée. A côté de ces utilisateurs enregistrés zélés (« zealots ») qui, motivés par l'acquisition d'une bonne réputation et/ou par l'adhésion à un projet communautaire, rendent compte publiquement de leur activité, ces bons samaritains anonymes (« good samaritans ») sont, a contrario, peu enclins à faire carrière ou à s'identifier à une communauté, mais participent bénévolement à cet effort intellectuel collectif.

Des auteurs occasionnels extérieurs à la communauté

Jimmy Wales ferait-il fausse route en répandant l'idée que Wikipédia n'est finalement guère différente d'une encyclopédie traditionnelle ? En appuyant son argumentaire sur le nombre de contributeurs qui cumulent le plus de modifications, il montre que 50% des modifications sont effectuées par 0,7% des utilisateurs (soit 524 personnes), alors que les 2% les plus actifs (1400 personnes) réalisent près de 75% de l'ensemble des modifications. Ce qui nous éloigne définitivement d'une distribution de type 80-20 (Pareto) et nous rapproche d'un schéma de type 90-9-1 :

- **90%** des utilisateurs sont des clandestins : ils lisent, observent mais ne contribuent pas ;
- **9%** sont des contributeurs occasionnels, dont les priorités sont extérieures à la communauté ;
- **1%** sont des contributeurs actifs, responsables de la grande majorité des contributions et démontrant pour certains un phénomène d'addiction.

Wikipédia serait-elle donc « écrite » par un petit millier de personnes ? Rien n'est moins sûr si l'on examine plus attentivement la nature de ces contributions, en particulier la proportion de texte ajoutée à chaque modification. Les résultats de l'expérience à laquelle s'est livré Swartz (2006) font largement écho à celle d'Anthony et al. (2005) citée précédemment. Il montre en effet que les utilisateurs qui apportent le contenu ne sont pas des contributeurs actifs, et pour une proportion non négligeable, ne se donnent même pas la peine de créer un identifiant. Ils sont intervenus en moyenne sur une dizaine d'articles, généralement liés à une même thématique. Se profile alors le modèle suivant, dont la similitude avec l'édition traditionnelle est évidente : un « outsider » édite l'article pour ajouter du texte, puis les « insiders » interviennent à plusieurs reprises pour faire des modifications mineures, reformater l'article ou bien l'affecter à une catégorie, une liste, etc. : actions que seules les personnes impliquées dans le projet peuvent accomplir. Alors que les « outsiders » fournissent le contenu, les « insiders » comptabilisent la majorité des révisions.

La présence de ces contributeurs situés à la périphérie du projet, anonymes ou non, demeure largement méconnue. La communauté Wikipédia, centrée sur elle-même, ne connaît pas ceux dont l'activité n'est pas publicisée. Les discours véhiculés mettent l'accent sur les chiffres, plutôt que sur la complexité de la dynamique éditoriale. Pour que l'objectif encyclopédique ne se voit pas minoré au bénéfice de l'activité gestionnaire, ces contributeurs occasionnels ne devraient-ils pas être mieux reconnus au sein de la communauté, même si tel n'est pas forcément leur souhait ? Du moins, leur existence permet-elle de relativiser les débats sur l'instauration d'une authentification préalable obligatoire...

Des collaborations à 2, 3 ou 4 paires de mains

L'écriture dans Wikipédia est incrémentale dans le sens où l'article se construit par petites touches. Selon Wagner (2006), les 20% d'existence d'un article comptent pour 35% et 36% de sa taille et pour 6% et 15% de ses éditions pour les articles créés respectivement en 2001 et 2002, avec un point de mesure fixé à mars 2004. Mais cette nature incrémentale correspond-elle à une juxtaposition de phrases et de paragraphes ou bien peut-on déceler des modèles d'interactions entre les contributeurs ?

Korfiatis et al. (2006) apportent un éclairage intéressant sur cette figure de l'auteur : ils s'appuient sur l'analyse des réseaux sociaux pour examiner les inter-relations entre articles et contributeurs. Tous les articles analysés (philosophie) répondent à un schéma rédactionnel décentralisé : l'écriture est distribuée entre de multiples contributeurs, démontrant par la même un fort degré de consensus. Elargie à la collection thématique, l'analyse montre là aussi un niveau de centralisation faible : autrement dit les auteurs participent à la rédaction de plusieurs articles. Ceux dont les interconnexions sont les plus fréquentes représentent les autorités les plus fortes, alors que ceux dont les interconnexions sont plus faibles se voient reconnaître un degré d'expertise moindre du fait que leurs contributions sont plus fréquemment contestées ou rejetées. L'étude montre donc que le processus rédactionnel s'organise autour de plusieurs groupuscules de contributeurs qui s'impliquent dans un ensemble d'articles représentatifs de leur communauté d'intérêt et dont les contributions dénotent une certaine forme d'expertise.

L'introduction d'algorithmes de réputation dans le MediaWiki peut-elle s'avérer pertinente pour mesurer la qualité des contributions de tel ou tel contributeur ? Si Korfiatis et al. (2006) semblent le suggérer, si Rozensweig (2006) en défend également le principe, un tel dispositif ne semble pas trouver d'écho dans la communauté aujourd'hui. Une métrique de la réputation, autrement dit une restauration de l'autorité individuelle ne s'inscrirait-elle pas en contradiction du projet démocratique et populaire ?

L'écriture à plusieurs mains : efficace ?

S'interroger sur la qualité de Wikipédia passe nécessairement à un moment ou un autre par la comparaison. Si l'édition traditionnelle n'échappe pas aux questions de fiabilité (Quid, Encarta, etc.), la suspicion est nécessairement plus aiguë s'agissant des contenus libres et collaboratifs. Plusieurs études permettent cependant de relativiser cette position.

Sur le fond tout d'abord, Wikipédia ne se distinguerait pas fondamentalement d'une encyclopédie classique. Les experts crédités par la revue *Nature* (Giles, 2005) relèvent une moyenne de 3 erreurs dans Britannica contre 4 dans Wikipédia sur les 42 articles scientifiques analysés. L'étude conduite par Chesney (2006) confirme ce résultat : les experts considèrent comme fiables les articles issus de leur domaine d'expertise, mais sont plus réservés sur ceux relevant d'une thématique qu'ils ne maîtrisent pas. De même, la comparaison entre Wikipédia, Encarta et American National Bibliography Online ne permet pas d'établir un différentiel significatif en termes factuels selon Rosensweig (2006).

A ce critère d'exactitude, s'ajoute celui de la formalisation linguistique et stylistique. Globalement meilleure dans Britannica, cette formalisation s'avère tout à fait proche dans les deux corpus, si l'on examine la fréquence des substantifs abstraits et des pronoms impersonnels d'une part, et la longueur moyenne des mots d'autre part (Elia, 2006). En comparant Wikipédia et Everything2²⁰, Emigh et Herring (2005) montrent également que le niveau de standardisation dans Wikipédia est comparable à celui de leur valeur témoin, l'Encyclopédie Columbia. Au contraire, les articles d'Everything2 font preuve d'un niveau de langue plus familier et sont stylistiquement moins homogènes. Les auteurs attribuent ce résultat au contrôle post-éditorial ouvertement collaboratif et à la ligne éditoriale explicitement formalisée de Wikipédia, remettant en question par là-même les systèmes basés sur la réputation.

Mais ces approches statistiques ne rendent que très imparfaitement compte de l'expression même. L'écriture à plusieurs mains, comme une juxtaposition de phrases ou de paragraphes rédigés par différents auteurs, ne se voit-elle pas technicisée au point de s'appauvrir sur les plans sémantique et stylistique ? Selon Rosenzweig (2006), cette difficulté constitue une des principales faiblesses de Wikipédia : inhérente au processus même, elle est également symptomatique d'une approche plus

²⁰ <http://www.everything2.com/> : plate-forme de production collaborative dont la régulation s'appuie sur un système de filtrage basé sur la popularité. Dans ce système à réputation, les articles sont signés par un seul auteur et les droits de publier affectés à ce même auteur sont techniquement dépendants des votes des autres utilisateurs de la plate-forme.

factuelle qu'analytique, contrairement à d'autres ouvrages de référence tels que l'American National Bibliography Online par exemple. Un écueil qui, selon lui, est aussi fortement lié à la concurrence manifeste entre la temporalité immédiate de la mise à jour et la temporalité historique et circulaire du savoir : quand l'actualité devient histoire, les frontières entre anecdotes et événements historiques sont poreuses.

Conclusion

Évaluer Wikipédia en termes d'échec ou de réussite nécessite de ne pas se limiter à une granularité minimale (celle des articles, voire même des phrases) et de considérer l'activité éditoriale dans sa dynamique globale et régionale. Si la croissance exponentielle des contributeurs ne renseigne guère sur la construction temporelle de la qualité, les indices d'une maturité éditoriale active sont nombreux. L'actualisation permanente des contenus s'accompagne très nettement d'une évolution constante des règles et des dispositifs techniques et humains au service de la régulation et de la validation éditoriale. Alors que les espaces du wiki dédiés à la gestion et à la dynamisation du projet se multiplient, les portails thématiques et wikiprojets se développent pour concentrer et organiser l'activité d'édition autour de besoins publicisés.

Parallèlement, les contributions se diversifient tout en se spécialisant, laissant émerger des formes d'expertise multiples, mais somme toute relativement traditionnelles. Les fonctions qui apparaissent sont significatives d'une tension croissante vers la professionnalisation. Nous serions grossièrement en présence de deux types de profils : des auteurs occasionnels, anonymes ou non, motivés par le projet éditorial mais peu enclins à faire carrière dans la communauté ; et des contributeurs fortement impliqués dans la dynamique communautaire qui construisent leur identité « professionnelle » autour de différentes tâches éditoriales. Les stratégies d'engagement des uns et des autres reposent sur une action critique participante dans un environnement qui favorise l'émulation des énergies créatrices personnelles.

Yochai Benkler (2002) suggérait déjà en 2002 que Wikipédia représentait un exemple type des modes « commons-based peer production », reposant sur l'exploitation de l'intelligence et de la créativité humaine et sur le volontarisme à s'engager dans un effort intellectuel conjoint. Plus récemment, Kolbitsch et Maurer (2006) montrent que l'ensemble de ces services web 2.0 – dont Wikipédia – ne sont pas simplement liés à l'apparition de nouvelles technologies mais s'inscrivent dans une évolution culturelle fondamentale qui encourage l'explicitation des savoirs de chacun et favorise le développement de l'intelligence collective. Dans un monde incertain, le savoir se démocratise et naissent des formes hybrides qui relèguent la traditionnelle distinction entre savoirs savants et savoirs profanes.

Ainsi ce qui pouvait encore très récemment être qualifié de « technologies disruptives » se banalise pour investir progressivement mais irrémédiablement le « mainstream ». Ces transformations laissent apparaître, dans le cas de Wikipédia, les prémices d'un modèle éditorial où production libre et collaborative d'une part et édition traditionnelle d'autre part s'auto-fécondent. Ce sont les recherches à venir qui viendront confirmer ou non ce qui se dessine à travers les analyses présentées dans cette contribution. Si la grande majorité des recherches portent aujourd'hui sur la version anglaise, on peut espérer que dans un avenir proche des analyses plus systématiques sur d'autres versions viendront affiner ces résultats et rendre compte de la complexité des logiques éditoriales.

Bibliographie

Les références utilisées pour cette contribution, enrichies de résumés et d'annotations, sont consultables et téléchargeables sur la wikibibliographie du projet Encyclen :

http://wikindx.inrp.fr/biblio_encyclen/

Anthony Denise, Smith Sean W., Williamson Tim. 2005. *Explaining quality in internet collective goods : zealots and good samaritans in the case of Wikipedia*. Hanover : Dartmouth College.

Benkler Yochai. 2002. Coase's penguin, or, Linux and the nature of the firm, *The Yale Law Journal*, vol. 112, n°3.

Bryant Susan L., Forte Andrea, Bruckman Amy. 2005. Becoming wikipedian : transformation of participation in a collaborative online encyclopedia. *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*, New York, ACM Press.

Carr Nicholas G. 2006. *Deletionists, inclusionists and delusionists*. Consulté le 15/12/06, <http://www.routhtype.com/archives/2006/09/a_fork_in_wikip.php>.

Chesney Thomas. 2006. An empirical examination of Wikipedia's credibility, *First Monday*, vol. 11, n°11.

Duguid Paul. 2006. Limits of self-organization : peer production and "laws of quality", *First Monday*, vol. 11, n°10.

Elia Antonella. 2006. An analysis of Wikipedia digital writing. *Proceedings of the EACL 2006 Workshop on New Text - Wikis and blogs and other dynamic text sources*. Trento, Italy (avril 2006).

Emigh William, Herring Susan C. 2005. Collaborative authoring on the web : a genre analysis of online encyclopedias. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, Los Alamitos, IEEE Press.

Endrizzi Laure. 2006. L'édition de référence libre et collaborative : le cas de Wikipédia. *Dossiers de la VST*. Lyon : INRP. Consulté le 15/12/06, <<http://www.inrp.fr/vst/Dossiers/Wikipedia/sommaire.htm>>.

Giles Jim. 2005. Internet encyclopaedias go head to head, *Nature*, n°438.

Holloway Todd, Božievi Miran, Börner Katy. 2005. *Analyzing and visualizing the semantic coverage of Wikipedia and its authors*. Bloomington: Indiana University. Consulté le 15/12/06, <<http://arxiv.org/ftp/cs/papers/0512/0512085.pdf>>.

Kolbitsch Josef. 2006. *Authorship in Wikipedia*. Graz, Austria: Graz University of Technology. Consulté le 15/12/06, <http://www.kolbitsch.org/research/papers/2006-JCMC-Authorship_in_Wikipedia.pdf>.

Kolbitsch Josef, Maurer Hermann. 2006. The transformation of the web : how emerging communities shape the information we consume, *Journal of Universal Computer Science*, vol. 12, n°2.

Korfiatis Nikolaos, Poulos Marios, Bokos George. 2006. Evaluating authoritative sources using social networks : an insight from Wikipedia, *Online Information Review*, vol. 30, n°3.

Lawler Cormac. 2005. *Wikipedia as a learning community*. Mémoire de Master, Manchester, University of Manchester.

Levrel Julien. 2005. *Développement des contenus libres sur internet – Wikipedia : approche de l'organisation en ligne et motivations des contributeurs*. France Telecom R&D TECH/SUSI.

Reagle Joseph. 2005. Case of mutual aid : Wikipedia, politeness, and perspective taking. *Proceedings of Wikimania 2005*. Frankfurt, Germany (août 2005).

Riehle Dirk. 2006. How and why Wikipedia works : an interview with Angela Beesley, Elisabeth Bauer, and Kizu Naoko. *Proceedings of the 2006 international symposium on Wikis*, New York, ACM Press.

Rosenzweig Roy. 2006. Can history be open source? : Wikipedia and the future of the past, *Journal of American History*, vol. 93, n°1.

Sanger Larry. 2005. The Early History of Nupedia and Wikipedia : A Memoir. In Chris DiBona,

Mark Stone, Danese Cooper. *Open Sources 2.0 : The Continuing Evolution*. O'Reilly.

Sanger Larry. 2006. Why make room for experts in web 2.0? *SDForum : the business of new media*. Santa Clara (octobre 2006).

Stvilia Besiki, Twidale Michael B., Gasser Les, Smith Linda C. 2005. *Information quality discussions in Wikipedia*. Champaign: University of Illinois at Urbana-Champaign. Consulté le 15/12/06, <<http://www.isrl.uiuc.edu/~stvilia/papers/qualWiki.pdf>>.

Swartz Aaron. 2006. *Who writes Wikipedia ?* Consulté le 15/12/06, <<http://www.aaronsw.com/weblog/howwriteswikipedia>>.

Viégas Fernanda B., Wattenberg Martin, Dave Kushal. 2004. Studying cooperation and conflict between authors with history flow visualizations. *Proceedings of the SIGCHI conference on Human factors in computing systems*, New York, ACM Press.

Viégas Fernanda B., Wattenberg Martin, Kriss Jesse, van Ham Frank. 2007. Talk before you type : coordination in Wikipedia. *Proceedings of the 40th Hawaiian International Conference of Systems Sciences*. Big Island, Hawaii (janvier 2007).

Voss Jakob. 2005. Measuring Wikipedia. *Proceedings of the 10th International Conference of the International Society for Scientometrics and Informetrics*. Stockholm, Sweden (juillet 2005).

Wagner Christian. 2006. Breaking the knowledge acquisition bottleneck through conversational knowledge management, *Information Resources Management Journal*, vol. 19, n°1.