L’édition de référence libre et collaborative: le cas de Wikipedia
Laure Endrizzi

To cite this version:
Laure Endrizzi. L’édition de référence libre et collaborative: le cas de Wikipedia. équipe LIRE - université Lyon 2. 2006, pp.1-32. sic_00473760

HAL Id: sic_00473760
https://archivesic.ccsd.cnrs.fr/sic_00473760
Submitted on 16 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
L’ÉDITION DE RÉFÉRENCE LIBRE ET COLLABORATIVE

Le cas de Wikipedia

Avril 2006
Les dossiers de la veille

Les *Dossiers de la Veille* ont pour vocation de présenter un état de l’art de la recherche sur une problématique, choisie et traitée à partir de références bibliographiques françaises et internationales.

**Déjà parus**

- Pratiques enseignantes (2006)
- Standards, compétences de base et socle commun (2005)
- L’enseignement supérieur sous le regard des chercheurs (2005)
L'édition de référence libre et collaborative : le cas de Wikipedia

Sommaire

Préambule .................................................................................................................................................. 4

Introduction ............................................................................................................................................... 5

1. Wiki et encyclopédie .................................................................................................................................. 6
   1.1. Wikis à vocation encyclopédique .................................................................................................... 6
   1.2. De Nupedia à Wikipedia .................................................................................................................. 6
   1.3. Principales critiques .......................................................................................................................... 7

2. Vers un savoir multiforme, ouvert et actuel ? .......................................................................................... 8
   2.1. Des contenus multilingues et traduisibles ....................................................................................... 8
   2.2. Architecture et catégorisation ......................................................................................................... 9
      2.2.1. Une architecture différenciée selon les univers linguistico-culturels ..................................... 9
      2.2.2. Une catégorisation libre et peu contrôlée .............................................................................. 9
   2.3. Des contenus libres de droits et réutilisables ................................................................................ 10
      2.3.1. Une instrumentisation ascendante du droit ......................................................................... 10
      2.3.2. Des modes d’accès hybrides et centrifuges ....................................................................... 10
   2.4. Des contenus actualisés en permanence ....................................................................................... 11
      2.4.1. Crédibilité : une question de stabilité ? ............................................................................ 11
      2.4.2. Des versions stabilisées en cours .................................................................................... 12
      2.4.3. Entre journalisme et encyclopédisme ? ........................................................................... 12

3. Vers un savoir co-construit, démocratique et perfectible ? .................................................................... 13
   3.1. Quelques dispositifs pour l’auto-gestion et l’auto-régulation ....................................................... 13
      3.1.1. Dispositifs transversaux pour la gouvernance globale ......................................................... 13
      3.1.2. Dispositifs techniques pour l’activité éditoriale .................................................................. 14
   3.2. Processus de stabilisation des articles et gestion des conflits .................................................... 15
      3.2.1. Coopération et gestion des conflits .................................................................................. 15
      3.2.2. Lutte contre le vandalisme ................................................................................................. 16
   3.3. Fiabilité et qualité ......................................................................................................................... 16
      3.3.1. Tests et mesures quantitatives ........................................................................................... 16
      3.3.1. Mesures qualitatives et construction temporelle ............................................................... 17
   3.4. Dialectique wiki / contenus ........................................................................................................... 17
      3.4.1. Le Bazaar style : une approche décentralisée efficace ....................................................... 17
      3.4.2. Les systèmes à réputation remis en question ...................................................................... 18

4. Les Wikipédiens : statuts, rôles et profils ............................................................................................. 19
   4.1. Une hiérarchie de pouvoirs techniques ....................................................................................... 19
   4.2. Des architectes aux profils diversifiés ......................................................................................... 19
   4.3. Des écoles de pensée qui s’affrontent ......................................................................................... 20
   4.4. Acculturation et création de sens .............................................................................................. 20
   4.5. Contributions périphériques ou anonymes ................................................................................ 21

5. Vers une organisation apprenante ? ...................................................................................................... 22
   5.1. Revue de projets pédagogiques ................................................................................................... 22
   5.2. Wikipedia comme organisation apprenante ? ........................................................................... 22
      5.2.1. Retours d’expérience sur deux projets pédagogiques ....................................................... 22
      5.2.2. Action critique et réflexivité explicite ............................................................................. 23

Conclusion .................................................................................................................................................. 24

Bibliographie ............................................................................................................................................ 25
   Recherches .......................................................................................................................................... 25
   Débats et réflexions ........................................................................................................................... 26

Sitographie .................................................................................................................................................. 28

La veille scientifique et technologique de l’INRP a conçu une base de connaissances agrégant plus de 300 références bibliographiques sur les encyclopédies numériques.

La bibliographie aborde des thématiques plus larges que celles traitées dans cette synthèse. Elle comporte les entrées suivantes :

- Fondements du savoir encyclopédique
- Informatisation et mutations technologiques
- Interfaces et modes de consultation
- Usages et pratiques de consultation

En complément, la base propose des références d’articles de débats et de réflexions, un ensemble de pages méta issues de Wikipedia, une liste de bibliographies en ligne et une sélection d’encyclopédies en ligne.

Cette bibliographie est mise à disposition via un outil wiki dédié à la gestion de références bibliographiques : Wikindx. Si vous souhaitez aller au-delà de la simple consultation, n’hésitez pas à prendre contact avec l’équipe de la VST.

L’année 2005 a été particulièrement riche en débats et controverses sur l’encyclopédie libre Wikipedia. Alors que l’attention des médias grands publics et les réactions des traditionnels médiateurs du savoir restent le plus souvent mobilisées par des affaires de vandalisme et plus généralement par des problèmes de fiabilité et de qualité, les usages ne cessent de se développer. Le nombre total d’articles sur l’ensemble des versions est passé de 1 million 400.000 à 3 millions 400.000 en l’espace d’un an. Cette croissance exponentielle des contenus s’accompagne d’une augmentation forte de la fréquentation, plaçant l’encyclopédie parmi les 25 sites les plus visités au monde selon le baromètre Alexa avec un trafic très largement supérieur à celui d’autres sites de référence tels que Britannica par exemple.


Nous nous attacherons dans cette synthèse à mettre en perspective les premières recherches sur Wikipedia avec les principales prises de position relayées par les médias traditionnels et l’internet.

Le lecteur ne s’étonnera pas de voir se côtoyer des travaux de chercheurs – quelquefois jeunes et méconnus – avec quelques articles de la presse généraliste ou spécialisée mais aussi avec des réactions et analyses d’acteurs de l’édition de référence (producteurs et prescripteurs). Ces dernières références seront matérialisées dans le texte par des crochets et un astérisque : [Auteur, année]* ; elles renverront à la partie « débats et réflexions » de la bibliographie.
Introduction

La « wikification » d’un corpus encyclopédique n’est pas une idée nouvelle. Linda Smith [Smith, 1989], chercheur à la Graduate School of Library and Information Science (University of Illinois, Urbana – Champaign), reprenant les prédictions de Vannevar Bush dans son article *As we may think* publié en 1945, anticipait déjà l’avènement de « wholly new forms of encyclopedias » qui dans un environnement hypertextuel démultiplieraient les accès au savoir, associeraient auteur et lecteur dans une dynamique de production commune et pointeraient directement sur des références externes. Ce qui en revanche est nouveau avec Wikipedia, ce sont ses barrières éditoriales relativement faibles, sa taille illimitée, sa rapidité dans l’actualisation des connaissances, son développement multilingue et sa relative simplicité technique [Stvilia et al., 2005].


1. Wiki et encyclopédie

Pour mieux appréhender les caractéristiques de ces nouveaux contenus encyclopédiques, intéressons-nous d’abord aux expériences de wikis encyclopédiques, au contexte présidant à la naissance de Wikipedia, et aux principales critiques dont Wikipedia fait l’objet.

1.1. Wikis à vocation encyclopédique

Le wiki est un site dynamique qui permet à toute personne disposant d’une connexion Internet d’en modifier les pages à volonté via son navigateur Web et à l’aide d’une syntaxe relativement simple. Le concept, théorisé par Ward Cunningham, a été implémenté pour la première fois en 1995. Aujourd’hui, il existe plusieurs dizaines de moteurs wikis aux fonctionnalités plus ou moins distinctives. A des degrés divers, tous prétendent des environnements collaboratifs où contenus, méta-informations et outils de communications sont intimentement corrélés. Tous favorisent l’autopublication collective et la structuration a posteriori des contenus. Ce sont les relations dialectiques entre les outils (techniques) et les méthodes (humaines) qui sont à l’origine des principales différences, comme nous le verrons dans le cas des systèmes à réputation.


Mais la production collaborative de contenus libres avec de tels outils ne se limite pas aux activités de cette fondation. S’agissant de contenus encyclopédiques, plusieurs expériences sont en cours, parmi lesquelles :

- l’Encyclopédia libre Universal en Español (castillan) ;
- le projet multilingue d’initiative universitaire Jurispedia consacré aux sciences juridiques et politiques ;
- l’encyclopédie PlanetMath pour les connaissances mathématiques ;
- les sites anglophones Everything2 et h2g2 et le site suédois Susning, qui se présentent plutôt comme des dictionnaires encyclopédiques et dont les outils d’édition offrent des paramétrages différents du Mediawiki utilisé par la fondation Wikimedia.

Ce mouvement s’étend par ailleurs à des sites affichant une idéologie alternative plus marquée : née en juillet 2003 d’une scission avec Wikipedia (fork), l’encyclopédie Wikinfo oppose à la neutralité de point de vue de Wikipedia (NPOV) une ligne éditoriale fondée sur le sympathetic point of view (SPOV) ; l’encyclopédie aux contenus libres Uncyclopedia née en janvier 2005, se présente comme une collection d’articles satiriques, voire absurdes parodiant Wikipedia ; le centre pour les médias et la démocratie Disinfopedia propose une collection d’articles d’actualité. Des articles dédiés à ces autres formes « encyclopédiques » peuvent par ailleurs être consultés dans la Wikipedia anglaise.

1.2. De Nupedia à Wikipedia


C’est la lourdeur de la chaîne éditoriale qui a eu raison de Nupedia et a conduit à sa fermeture définitive en septembre 2003 après plus d’un an d’inactivité, alors que 24 articles avaient été intégrés et 74 autres étaient en développement. Parallèlement, Wales et Sanger avaient lancé Wikipedia (janvier 2001 pour la version anglaise), avec l’objectif d’expérimenter un outil plus flexible qui faciliterait la production collaborative et décentralisée des articles, mais sa notoriété exponentielle les a amenés à recentrer leurs efforts sur ce nouveau projet au détriment de Nupedia.

En avril dernier, Sanger [Sanger, 2005a et b]* rappelle qu’à l’origine Wikipedia, née avec un.com, ne se positionnait pas en faveur d’une démocratisation totale des savoirs. Lancée provisoirement sans règles précises pour mieux appréhender la flexibilité de wiki, l’encyclopédie libre a évolué les premiers mois dans « a good-natured anarchy, a sort Rousseauian state of digital nature » gouvernés par un consensus implicite entre quelques contributeurs en provenance de Nupedia : un noyau de 200 personnes environ, déjà sensibilisés au projet et fortement engagées dans cette dynamique collective.

Dès les premiers mois, les principes de base ont cependant été édictés et trois constituent encore aujourd’hui les pivots inamovibles du projet (cf. Policies and guidelines ou Recommandations et règles à suivre) :

- la neutralité de point de vue (neutral point of view ou NPOV);
- le respect du copyright et la licence libre GNU GPL (general free documentation licence);
- le respect des contributeurs dans le processus de rédaction collective (wikipetiquette).

La réussite de Wikipedia repose selon Larry Sanger [Sanger, 2005b]* sur quelques facteurs soit empruntés au mouvement open source ou à la culture wiki soit plus spécifiques au projet même, dont il résume ainsi la combinaison :

« The focus on the encyclopedia provided the task and the open content license provided a natural motivation : people work hard if they believe they are teaching the world stuff. Openness and ease of editing made it easy for new people to join in and get to work. Collaboration helped move work forward quickly and efficiently, and posting unedited drafts made collaboration possible. The fact that we started with a core of good people from Nupedia meant that the project could develop a functional, cooperative community. Neutrality made it easy for people to work together with relatively little conflict. And the Google effect provided a steady supply of “fresh blood” who in turn supplied increasing amounts of content. »

Selon lui, c’est l’arrivée de la deuxième vague de contributeurs qui a cristallisé cette anarchie initiale et qui a rendu le projet ingérable et toute responsabilité éditoriale illusoire. S’étant retiré peu après que la société Bomis ait décidé de ne plus le rémunérer pour sa fonction d’éditeur en chef (courant 2002), Larry Sanger refuse de critiquer pour qui la gouvernance de Wikipedia demeure un échec.

À la lumière de son expérience personnelle, il regrette que des procédures de contrôle plus strictes n’aient pas été déterminées en amont : par exemple établir des règles non négociables et rédiger une charte soumise à l’adhésion des participants, définir une autorité qui garantisse le respect des règles et principes et qui sanctionne les comportements perturbateurs sans état d’âme, et pour finir

---


intégrer la notion d’expertise et de méritocratie en prévoyant des rôles spécifiques pour les experts dans le processus éditorial. La rumeur selon laquelle il pourrait lancer un projet alternatif n’en est d’ailleurs plus une : dans une lettre ouverte à la communauté scientifique en janvier 2006 [Sanger, 2006]*, il lance un appel à mobilisation auprès de la communauté scientifique pour la création d’une nouvelle encyclopédie fondée sur la promotion d’un Internet d’experts.

1.3. Principales critiques


En résumé, les principales critiques à l’égard d’un tel outil de publication sont focalisées sur les points suivants :

- les contributeurs sont au mieux des amateurs, au pire des perturbateurs, des vandales ou des fanatiques
- quand ils ne se dissimulent pas derrière un ou plusieurs pseudonymes, les contributeurs sont des anonymes
- les contributeurs sincères sont découragés par la nécessité de défendre leurs textes contre d’autres participants ignorants ou malveillants
- les contributeurs, non experts, sont illégitimes
- les contributeurs ne sont juridiquement pas responsables de leurs contributions
- hormis ceux traitant d’informatique, les articles sont au mieux de qualité médiocre, au pire de simples ébauches
- certains thèmes ou sujets controversés sont décrédibilisés par des prises de positions partisanes (biographies, événements historiques, mouvements politiques, religion, etc.)
- les sujets d’actualité sont sur-représentés au détriment d’un savoir établi
- les sources sont rarement indiquées, le contenu n’est pas vérifiable
- le non respect du copyright (plagiat) et le non respect de la vie privée sont des pratiques courantes
- les articles ne sont pas stables et peuvent être supprimés ou vandalisés à tout moment
- le manque de fiabilité et de stabilité des articles les rend inexploitables
- le qualificatif « encyclopédie » est scientifiquement inapproprié
- la régulation collective est absurde et illusoire
- l’activité éditoriale et plus généralement la structuration des contenus manquent de transparence
- la croissance exponentielle des articles et des participants rend illusoire toute tentative d’homogénéisation et de stabilisation
- le principe de la neutralité de point de vue est en fait éminemment politique
- la vérité encyclopédique est « votée » ; quelle représentativité pour les votants ?
- le fondateur de Wikipedia est un personnage immoral qui exerce une influence dictatoriale
- la représentativité des univers culturels des locuteurs d’une même langue est biaisée
- etc.

Cette synthèse ne prétend pas répondre à toutes les questions soulevées par cette rapide revue de critiques. Ce sont les recherches exploitées qui nous permettront d’approfondir la réflexion sur certaines d’entre elles.
Selon Bertrand Legendre [Legendre, 2000], chercheur en sciences de la communication à l’université Paris 13, le développement de l’édition numérique constitue à plusieurs titres une rupture dans les traditions et habitudes culturelles dont l’objet encyclopédique est emblématique. Cette mutation se manifeste par un « élargissement » constant des formes qui substitue à une somme de savoirs autosuffisants et complémentaires répondant à une logique circulaire, un ensemble multiforme « construit par simple juxtaposition » d’informations et de sources, qui affecte le dispositif de légitimité culturelle fourni par l’édition traditionnelle.

En 2000, la Webencyclo des éditions Atlas paraît représentatif de cette évolution : d’une part, le contenu initial de l’encyclopédie (des articles relativement brefs) est enrichi par des dossiers thématiques conséquents, et d’autre part les lecteurs sont encouragés à contribuer au développement de cette offre des dossiers, soit en proposant de nouveaux thèmes, soit en apportant directement des contenus.

La fonction éditoriale se déplace vers une fonction d’accumulation qui relève au second plan le rôle de mise en cohérence et de construction circulaire des contenus encyclopédiques. Tout se passe comme si le passage de l’espace clos du livre aux espaces illimités du réseau soumet la fonction éditoriale à la nécessité de combler ces ouvertures, notamment par l’intégration parfois quasi « promotionnelle » d’éléments multimédias, et par un glissement progressif vers le ludo-éducatif.


Avec Wikipedia, cette externalisation est poussée à son paroxysme puisque tous les contributeurs sont au départ comme « extérieurs » au projet éditorial. Comme le souligne notamment Andrea Cifollili [Cifollili, 2003], doctorant en économie à l’université d’Ancona en Italie, le coût de la transaction dans l’édition d’un article de Wikipedia est relativement bas, comparé à celui d’une transaction similaire dans un environnement peer-review. Si l’édition numérique peut en quelque sorte s’affranchir des contraintes de production liées à la forme papier, Wikipedia franchit un pas supplémentaire en incorporant littéralement cette absence de limites. Alors qu’Encarta annonce 36.000 articles dans son édition 2006, Universals compte aujourd’hui 30.000 et Britannica 75.000, la Wikipedia anglaise revendique plus de 935.000 entrées et sa version française près de 230.000 en janvier 2006, après à peine 5 ans d’existence (et l’on se souvient encore peut être du communiqué d’avril 2005 qui annonnçait le franchissement des 100.000 articles en français).

2. Vers un savoir multiforme, ouvert et actuel ?

2.1. Des contenus multilingues et traduisibles

2.1.1. Un multilinguisme inédit

Wikipedia se décline dans 200 langues différentes, incluant des langues régionales telles que le basque ou le breton et des langues mortes telles que le latin ou le sanskrit par exemple. 8 versions régionales comptent plus de 100 000 articles, 27 plus de 10 000 et 48 plus de 1 000. Sur un total de 1 million 400 000 articles en janvier 2005, la proportion d’articles en langue anglaise représentait 33%. Au moment où nous écrivons, la Wikipedia française tient la 3ème place avec 230.000 articles derrière les Wikipédias anglaise (935.000) et allemande (346.000), alors que le nombre total d’articles sur l’ensemble des versions atteint presque 3 millions 400.000 (cf. dans la catégorie Statistics : Multilingual statistics ou les Wikistats par pays). Chaque version régionale gère ses propres contributeurs et fonctionne selon des modes organisationnels variés dans le respect des principes fondateurs et des règles générales d’édition.

La coordination entre les langues s’avère plus ou moins structurée selon les versions (cf. multilingual coordination and coordination multilingues) et vise à promouvoir l’établissement de renvois multilingues entre les articles et la traduction d’articles de qualité.

Par exemple, à partir de l’article New York de la Wikipedia française, on peut accéder à l’article équivalent dans une vingtaine de langues. Les liens entre les articles sont appelés interlanguage links ou liens interlangues et répondent à une syntaxe propre à la gestion des liens InterWiki utilisés entre les différents projets de la fondation Wikimedia.

Les contenus, réutilisables dans le cadre de la licence GFDL (general free documentation license) donc traduisibles, font l’objet d’appels à traduction réguliers sur les pages d’accueil des communautés (cf. community portal et communauté). Alors que le projet traduction de la Wikipedia française est en gestation au moment de la rédaction, la Wikipedia anglaise dispose déjà d’un département dédié à la traduction des projets linguistiques bilatéraux pour faciliter la sélection d’articles et opérationnaliser la collaboration (cf. par exemple la page French collaboration project).

Aucune statistique à notre connaissance ne permet de rendre compte de l’ampleur de cette activité de traduction. Cette tendance, dynamisée par la croissance de la Wikipedia, doit-elle être considérée préjudiciable au développement autonome des différents projets linguistiques ? La traduction favorise-t-elle perceptuellement l’homogénéisation, voire l’uniformisation des contenus au bénéfice d’une langue anglaise et d’une culture anglo-saxonne hégémonique ? Ou bien au contraire, constitue-t-elle un levier pour le développement de projets encore jeunes où le nombre de bénévoles est faible ?

Indépendamment de la traduction, le traitement d’un même sujet par différents univers linguistiques et culturels introduit de facto une vision du monde qui relève la représentation d’une connaissance universelle et objective. La seule encyclopédie actuellement multilingue, Encarta, a été justement pointé du doigt, comme produit commercial qui annihilait tout contexte socio-culturel, [Hay, 2004].

Aujourd’hui la tendance est en quelque sorte inversée, puisque nombre de détracteurs de Wikipedia se focalisent sur ces biais culturels. Comment un projet multilingue, multiculturel et multiforme est-il conçu et construit ? Quelle est l’importance de ces visions du monde dans Wikipedia ? Dans quelle mesure une version linguistique donnée est-elle représentative des différents contextes de production nationaux partageant cette même langue ? Dans quelle mesure ces mêmes contextes introduisent-ils des parts pris dans le traitement des sujets ?

2.1.2. Cartographie et biais culturels

Francesco Bellomi et Roberto Bonato [Bellomi et al., 2005], tous deux doctorants en informatique à l’université de Vérone (Italie), se sont intéressés à cette question des biais culturels en appuyant leur étude sur l’analyse de réseau (network analysis). Cette méthode, issue de la théorie des graphes, vise à décrire les propriétés des entités interconnectées au moyen d’algorithmes mathématiques et à mesurer la pertinence des contenus (content relevance ranking). L’application des algorithmes Hyperlink-induced topic selection (HITS) et PageRank (PR) au corpus de la Wikipedia anglaise apporte un éclairage tout à fait intéressant sur la macro-structure de l’encyclopédie.

Pour mieux comprendre, le PR est l’indice de popularité d’une page, calculé avec le nombre de pages qui pointent vers elle : c’est l’algorithme banalisé par Google ; le HITS identifie les pages qui font autorité (authority) c’est-à-dire qui sont les plus pointées par des
pages qui concentrent les renvois (hubs). Si ces deux algorithmes mesurent le nombre de citations d’une page, les autorités sont des pages de référence bénéficiant d’une structuration thématique forte, contrairement aux pages populaires. Les résultats montrent que Wikipedia forme un seul graphe de pages interconnectées et qu’il n’existe donc pas d’îlots séparés du reste du corpus. Avec le HITS, il apparaît que les catégories les plus structurantes et donc les plus structurées dans Wikipedia sont celles relatives à l’espace géopolitique et au temps (périodes et événements historiques) et dans une moindre mesure, celles concernant les personnes, ainsi que les noms courants tels music, philosophy ou religion. Le PR, en revanche, révèle une écrasante domination des concepts étroitement liés à la religion dans les 300 premiers résultats : les pages les plus populaires, et donc le plus souvent citées sans pour autant être considérées comme des pages de référence, sont donc clairement orientées sur un thematique singulière, celle de la religion.

Dans une seconde étape, la comparaison du HITS et du PR pour certaines classes thématiques permet d’identifier une autre forme de biais internes à ces mêmes classes. Le classement des pages populaires s’avère plus varié et global pour les classes « pays et villes », « événements historiques » et « personnalités », alors que celui des pages de référence révèle un ancrage plus fort de ces thèmes dans l’actualité et un recentrage sur les États-Unis. À l’inverse, les trois premières entrées dans la classe « noms communs » montrent que les biais sont fournis par l’indice de popularité : pope, god et priest.

Si ces premiers résultats confirment l’existence de biais culturels, historiques et politiques dans la Wikipedia anglaise, ils ont le mérite de faire tomber quelques idées reçues parfois tenaces (sur-représentation des articles informatiques notamment) et surtout de définir des outils de mesure pour élargir les investigations tant en synchronie, qu’en diachronie. Ils permettent également de mettre en lumière les déficits pour mieux mobiliser les membres de Wikipedia (cf. projet Countering systemic bias dans la Wikipedia anglaise).

2.2. Architecture et catégorisation

2.2.1. Une architecture différenciée selon les univers linguistico-culturels

Nous l’avons évoqué et nous y reviendrons plus en détail dans les troisième et quatrième parties de cette synthèse, le wiki propose dans une même interface des contenus encyclopédiques, des métainformations sur l’activité éditoriale et des outils de communication. La comparaison de l’importance de ces différents espaces dans les Wikipédias allemande et japonaise, respectivement 2ème et 4ème en terme de nombre d’articles, et dans les Wikipédias danoise et croate, de taille moyenne, apporte un éclairage nouveau sur la manière dont les contributeurs d’un même univers linguistique s’approprient le wiki et posent les fondations de l’encyclopédie. Avec une extraction complète des données de ces quatre bases, datée de janvier 2005, Jakob Voss [Voss, 2005a], doctorant en sciences de l’information à l’université Humboldt à Berlin, montre que les pages correspondant aux articles encyclopédiques même représentent de 60% pour la version allemande à 80% pour la version croate. Autrement dit, les 20 à 40% des contenus restants sont composés par des contenus multimédias (images, documents audio et vidéo illustrant les textes encylopédiques), des pages de discussion attachées aux articles encyclopédiques, des pages de catégories offrant une navigation thématique au sein de l’encyclopédie, des pages de modèles pour faciliter la création d’articles homogènes (menu de navigation dans un thème donné ou plan de présentation pour un type d’articles donné), des pages méta dédiées à la gestion même de Wikipedia et des pages personnelles pour les usagers enregistrés qui souhaitent rendre compte publiquement de leur activité au sein du projet.

Plus la version étudiée est jeune, plus la part des articles encyclopédiques est importante, l’article constituant la cellule de base autour de laquelle d’autres cellules viennent se greffer. Cumulée aux contenus multimédias et aux pages de discussion, cette proportion atteint globalement 90% de l’ensemble et 95% pour la Wikipédia croate. C’est la version allemande qui toutefois comprend le plus de pages multimédias (18%) et de pages de discussion (11%). Alors que les pages modèles représentent moins de 1% dans les quatre versions, les pages de catégories s’échelonnent de la manière suivante : 1% (langue croate), 2% (langue japonaise), 3% (langue allemande) et 5% (langue danoise), démontrant des niveaux de structuration différents. Le nombre de pages personnelles, relativement faible dans les versions de taille moyenne (2%), atteint 7% dans la Wikipédia allemande et 6% dans son homologue japonaise, alors même que le nombre de contributeurs de langue japonaise est 2 fois plus important, résultat qui laisse supposer que ces contributeurs sont plus réservés, voire préfèrent garder l’anonymat.

2.2.2. Une catégorisation libre et peu contrôlée

Les articles dans Wikipedia sont intégrés à des catégories hiérarchiques réparties entre une catégorie principale et des catégories racines, intermédiaires ou terminales dont la gestion s’opère par des relations d’inclusion (cf. Categorization ou Catégorie). La construction de cette catégorisation s’apparente à la folksonomie (apposition libre de tags ou d’étiquettes par les utilisateurs) dans la mesure où chaque contributeur peut indexer les articles et définir de nouvelles catégories mais s’en distingue par son fonctionnement arborescent.

Le fait que ni la catégorisation ni le nommage des articles ne bénéficient d’outils pour le contrôle d’intégrité génère trois principales difficultés : la désambiguïsation et la gestion des doublons, une structuration floue des relations entre catégories ou articles et une indexation plus ou moins aléatoire contrairement aux systèmes documentaires traditionnels. Les pages de redirection et de désambiguïsation créées par les contributeurs, gérant respectivement la synonymie et l’homonymie, n’apportent pas de réponse satisfaisante à ces difficultés.

Outre les pages qui concentrent les liens internes sur un thème ou une catégorie sous forme de listes alphabétiques ou structurées (lists guideline et list of topic lists dans la version anglaise), les portails ou plages agrégent des contenus thématiques au sein d’une même page pour offrir un mode alternatif de navigation représenté par des entrées macro ou micro qui favorisent la transversalité, contrairement aux catégories hiérarchiques. Ils apportent un début de réponse quant à la gestion thématique des articles, toutefois insuffisant. Le manque de visibilité de ces dispositifs et de leur articulation crée une confusion pour appréhender le traitement d’un thème donné, au-delà de l’unité de l’article.

De même au niveau des articles, si la distribution des liens hypertextes internes et externes présente une progression parallèle et proportionnelle à la croissance des articles et des contributions, l’évolution des renvois entre articles s’inscrit dans un processus complexe et peu homogène. J. Voss montre par exemple que pointer vers un article inexistant est une pratique courante dans la Wikipédia allemande, pour attirer l’attention sur un besoin : plus de 10.000 articles souhaités (wished articles) sont ainsi signalés par un lien hypertexte vide (red) [Voss, 2005a].

Plusieurs recherches aujourd’hui s’attachent à opérer des rapprochements avec les travaux issus du traitement automatique des langues et du Web sémantique pour optimiser cette structuration autogérée : [Harth et al., 2005], [Ruiz-Casado et al., 2005a et b], [Krötzsch et al., 2005], [Voss, 2005b] par exemple. Plus proche de nous, les travaux menés dans le cadre du projet Autograph (auto-organisation et visualisation des graphes, soutenu financièrement par l’ANR, comportent un volet sur Wikipedia et rassemblent des chercheurs du LIMSI, France Telecom R&D, ENST, LIAFA, INRIA Futurs, FING].
2.3. Des contenus libres de droits et réutilisables

2.3.1. Une instrumentalisation ascendante du droit

L’exemple des éditions Atlas cité précédemment mérite de retenir notre attention car, avant même le lancement de Wikipedia, la Webencyclo a donc été la première encyclopédie francophone à proposer l’ensemble de ses contenus sur le Web (cf. communiqué de décembre 1999), moyennant une inscription sur le site. Le terme « moyennant » n’est pas anodin comme l’explique cet article du Journal du Net (juillet 2002), puisque le modèle économique sous-jacent consistait à commercialiser la base de courriels ainsi constituée. L’appel à des contributions externes (cf. communiqué de mai 2000), a soulevé par ailleurs d’autres interrogations, les contenus restant placés sous le copyright des éditions Atlas. Les contributeurs étaient tenus de céder leurs droits à l’éditeur, tout en endossant la responsabilité pénale et civile à l’égard des tiers (et de Webencyclo) en cas de litige. Une stratégie commerciale qui s’est avérée finalement peu compatible avec les évolutions de Internet, le site ayant fermé ses portes « pour une durée indéterminée » le 30 octobre 2005 (courriel envoyé aux abonnés).

À côté de ces deux grands, de nouveaux moteurs de recherche orientés « référence » font leur apparition. La majorité de ces sites renvoient sur des sources totalement ou partiellement consultables en libre accès et proposent le plus souvent des espaces publicitaires ou des liens sponsorisés. La sélection des sources souffre parfois d’un manque de transparence et la consultation des articles s’opère le plus souvent au sein même de l’interface féédérée. Autrement dit, contrairement aux moteurs de recherche cités précédemment, ces outils ne génèrent pas systématiquement de trafic sur le site source.

2.3.2. Des modes d’accès hybrides et centrifuges

Si l’une des conséquences majeures de ces dispositions juridiques est qu’elle autorise les chercheurs à exploiter le corpus complet et mouvant des bases de données régionales, la licence GFDL constitue également un vecteur de diffusion essentiel pour démultiplier les points d’accès à l’encyclopédie.

Wikipedia indexée par les moteurs de recherche


Avec une requête de type define [subject], Google affiche une liste de définitions issues de Wikipedia, du CIA World Factbook et d’autres ouvrages de référence en ligne. Cette fonctionnalité semble opérationnelle en allemand, anglais, espagnol, français, italien, russe. Le service Google Answers s’appuie également sur ces mêmes ouvrages pour répondre aux questions des usagers.

La page Search engine statistics dans la Wikipedia anglaise permet de mieux se représenter ces relations avec les moteurs de recherche, Google notamment : elle fournit des données précises sur le PageRank de l’encyclopédie libre, ses citations dans les services Google News ou bien Google Answers et le nombre de pages indexées par le moteur. À côté de ces deux grands, de nouveaux moteurs de recherche orientés « référence » font leur apparition. La plupart des résultats avec Factbites proviennent de sites de référence et notamment d’encyclopédies en ligne, Wikipedia inclus. Les sources en accès libre côtoient celles accessibles sur abonnement via les grands fournisseurs en IST tels que Elsevier, Wiley, etc.

Clusty affiche également en priorité des pages issues de sites de référence en accès libre dont celles de Wikipedia. Un onglet spécifique permet par ailleurs d’interroger directement l’encyclopédie.

Des métamoteurs ou interfaces féédérées pour dictionnaires et encyclopédies

Avant même que les moteurs généralistes ne s’intéressent aux dictionnaires et encyclopédies, les métamoteurs ont commencé à fleurir et le nombre de services féédérés ne cesse de croître, essentiellement dans la sphère anglophone de Internet. La majorité de ces sites renvoient sur des sources totalement ou partiellement consultables en libre accès et proposent le plus souvent des espaces publicitaires ou des liens sponsorisés. La sélection des sources souffre parfois d’un manque de transparence et la consultation des articles s’opère le plus souvent au sein même de l’interface féédérée. Autrement dit, contrairement aux moteurs de recherche cités précédemment, ces outils ne génèrent pas systématiquement de trafic sur le site source.
L’édition de référence libre et collaborative : le cas de Wikipedia

Dès 1996, Onelook offre une interface de recherche spécialisée qui interroge aujourd’hui plus de 1000 dictionnaires, encyclopédies, et glossaires, très majoritairement en langue anglaise. Le site propose également des services de traduction et de recherche inversée pour explorer les relations sémantiques d’un mot. Une extension est disponible pour le navigateur Firefox.


Le site Threfeedictionary (Farlex, Inc.) fournit l’accès au dictionnaire American Heritage® Dictionary of the English Language, aux encyclopédies Columbia, Wikipedia et Hutchinson (articles complets réservés aux souscripteurs) et à d’autres ouvrages de référence plus spécialisés. Le site permet une navigation alphabétique et/ou sémantique (rebonds) et met à disposition des usagers et des webmestres des extensions et fils RSS pour l’interrogation à distance.

Comme pour Threfeedictionary, Encyclopedia.com permet d’accéder gratuitement aux contenus de la Columbia Encyclopedia, alors que celui de la Hutchinson est accessible sur abonnement. Le service, proposé par HighBeam™ Research, Inc., comprend aussi des dictionnaires, almanachs et thésaurus et pointe vers des sites de documents sur les diverses questions de bibliothèque numérique de la société. Les éditeurs et libraires traditionnels se lancent aussi dans le créneau. Ainsi le groupe Pearson Education Publishing agrège sur son site Infoliase l’accès à une encyclopédie (Columbia Encyclopedia), un dictionnaire, un atlas, un almanach, un thésaurus et un annuaire de sites pour les États-Unis, alors que le site Factmonster est un service similaire dédié à une cible junior. Du côté des libraires, c’est Bartleby qui fournit l’interface fédérée la plus développée, incluant des ouvrages spécialisés, des dictionnaires de citations, des thésaurus, etc.

Le site Xrefer se présente pour sa part comme une bibliothèque numérique de plus de 200 ouvrages de référence sélectionnés parmi les plus grands éditeurs internationaux et cible les bibliothèques : l’accès est réservé aux abonnés. Parmi les éditeurs signataires, on retrouve le Columbia Encyclopedia et très prochainement l’encyclopédie Britannica.

Dans la sphère francophone, le seul service similaire (gratuit) est fourni par DicoWeb qui permet de lancer des recherches successives dans quelques dictionnaires (unilingues, bilingues, synonymes, acronymes, etc.) et encyclopédies en ligne : ARTFL, Encarta, Wikipedia parmi les sources majeures.

Ce rapide tour d’horizon non exhaustif fournit quelques indices sur les stratégies commerciales offertes de certains éditeurs qui multiplient les points d’accès à leurs ouvrages : la Columbia Encyclopedia, les dictionnaires de Merriam Webster et ceux de Houghton Mifflin (American Heritage® Dictionnaires) par exemple. S’agissant de la Columbia Encyclopedia, notons toutefois que l’accès libre concerne des éditions anciennes (2001 ou 2003 selon les sites) et que des fonctionnalités de navigation telles que les renvois entre articles ne sont pas disponibles sur les serveurs externes.

**Des dictionnaires et encyclopédies intégrés au navigateur ou adaptés aux technologies mobiles**

Nous avons évoqué précédemment, les modes d’accès aux encyclopédies non seulement se multiplient mais se délocalisent en quelque sorte. Les extensions permettant d’interroger Wikipedia depuis le navigateur Firefox fleurissent (cf. liste multi-versions) ; ce sont des barres d’outils multi-sources telles Quero ou Yubnub ou bien encore des raccourcis de recherche à placer dans des favoris (celui de Geckozone ou bien Linky pour naviguer entre Amazon, Wikipedia et le catalogue virtuel des bibliothèques allemandes KVK). Les dictionnaires de Merriam Webster et de Cambridge mettent également à disposition des extensions pour Firefox (cf. liste des extensions dictionnaires et projet ConQuery sur le site Mozdev).

Les extensions contextuelles pour accéder à une entrée de l’encyclopédie au cours de la navigation se développent également : Hyperwords par exemple. De la même façon, le projet Dict sur Mozdev permet d’interroger le Collaborative International Dictionary of English mais peut être configuré avec n’importe quelle base de données. La Google Bar quant à elle offre désormais un traducteur mot à mot (affichage de la traduction d’un terme anglais) sur simple survol de la souris. À l’inverse les serveurs peuvent aussi être équipés d’outils linguistiques : Alexandria, conçu par Memodata et commercialisé par I-km, est un dictionnaire multilingue contextuel pour site Internet ou intranet et basé sur Wordnet.

La géolocalisation, technique qui permet de localiser sur une carte la position d’un objet, fournit également des vues intéressantes pour accéder autrement à Wikipedia : PlaceOpedia qui pointe sur les articles géographiques à partir d’une carte ou bien encore cette extension qui permet de localiser les contributeurs au Mediawiki par exemple.

Le site Wikiwax quant à lui aide l’internaute à identifier les entrées d’index de Wikipedia au fur et à mesure de la frappe, un peu à la manière de Google Suggest.


**2.4. Des contenus actualisés en permanence**

Avec Wikipedia, l’actualisation des contenus encyclopédiques, inscrite dans une temporalité beaucoup plus longue avec le format papier, remet en cause le savoir au bénéfice de l’information ? Ou bien renouvelle-t-elle le genre encyclopédique en reléguant son obsolescence ? Comment savoir si la version en cours est de meilleure qualité que la précédente ? Entre état et construction, comment s’y retrouver ?

**2.4.1. Crédibilité : une question de stabilité ?**

Avant même le lancement de Wikipedia, Bertrand Legendre [Legendre, 2000], estimait déjà que la transposition du papier au numérique soumettait l’édition traditionnelle au « risque de l’information » et que l’éditeur se devait de négocier « le passage d’une logique culturelle de la formation critique de la connaissance à une logique informationnelle » en privilégiant ainsi une « approche factuelle des contenus plutôt qu’une approche analytique ».

Selon Edward Feiten [cité dans Crawford, 2005], chercheur en informatique à l’université de Princeton, le processus éditorial s’apparente à un chemin aléatoire vers la qualité : « a random walk around some plateau of quality that may or may not be as good as...»
(or better than) a traditional encyclopedia, but is unlikely to be "the highest degree of accuracy" that Wikipedia’s advocates assert as a reachable goal.

Robert McHenri, ex-responsible éditorial chez Britannica [McHenri, 2004]*, estime quant à lui qu’une production collective remaniée en permanence ne peut pas tendre à la vérité encyclopédique à la manière d’une asymptote. La suspicion à l’égard de contenus non stabilisés est selon lui amplifiée par le fait que les utilisateurs n’ont aucune garantie sur l’identité des contributeurs.

Faut-il dès lors « stabiliser » les articles dans Wikipedia ? Si l’idée paraît à première vue peu compatible avec la philosophie même de l’encyclopédie libre, plusieurs voix, dont celle de la Fondation Internet nouvelle génération [Klein et al., 2006], sont favorables à l’instauration de modes collaboratifs dérivés du monde des logiciels libres, alors que parallèlement des éléments concrets nous autorisent à penser que Wikipedia s’achemine vers une standardisation progressive de ses contenus.

2.4.2. Des versions stabilisées en cours

Alain Caraco [GREMI, 2005], conservateur de bibliothèque, rappelle en effet que des réflexions sont en cours pour introduire des indications sur le statut de l’article (cf. Stable versions) à la manière des versions release utilisées pour les logiciels libres. Le déploiement d’une Wikipedia 1.0 (cf. Pushing 1.0) pourrait dès lors servir de filtre pour des éditions partielles stabilisées.


- commercialisation d’un cédérom et d’un dévédérom sous l’égide de l’association Wikimedia Deutschland en collaboration avec l’éditeur Directmedia Publishing GmbH
- mise à disposition de versions téléchargeables et enregistrables
- édition (papier) de fascicules thématiques appelés wikireaders

La 2e édition du dévédérom allemand est sortie en décembre 2005. Elle contient 300.000 articles et 100.000 images et est accompagnée d’un livre de 139 pages présentant Wikipedia, son histoire et son fonctionnement. Elle est vendue 9,90 euros et est notamment distribuée par Amazon.


2.4.3. Entre journalisme et encyclopédisme ?


Alors que la fondation Wikimedia lance Wikinews en décembre 2004, les professionnels de la presse s’intéressent également au wiki. Après l’échec de l’expérience Wikitorial du Los Angeles Times où les internautes étaient invités à participer aux éditoriaux du journal, c’est au tour du magazine Esquirre de publier un article « écrit à 150 mains » [Richard, 2005]* dont le brouillon truffé d’erreurs avait été soumis à la communauté Wikipedia.

Andrew Lih, chercheur au Journalism and Media Studies Centre à l’université de Hong Kong, montre que Wikipedia tend par ailleurs à être de plus en plus citée comme source dans la presse et que cette promotion indirecte n’est pas sans influence sur l’évolution même des contenus. Il établit également une relation forte entre le principe de la neutralité de point de vue (NPVO) de Wikipedia d’une part et le traitement des actualités dans la presse d’autre part : « sticking to the facts, attributing sources and maintaining balances » [Lih, 2004a].

Selon lui, les wikis, et en particulier celui de Wikipedia, sont des outils adaptés au développement du journalisme participatif (participatory journalism) [Lih, 2004b], notion popularisée par Dan Gillmor dans son ouvrage We the media [Gillmor, 2006]. La participation du public au processus de rationalisation des contenus d’information ouvre de véritables perspectives au journalisme au sein d’une nouvelle écologie des médias. Dans cette configuration, Wikipedia est perçue comme un espace idéal qui permettrait de combler le fossé de la connaissance entre les médias de masse et les publications historiques scientifiques, pour devenir en quelque sorte une mémoire de l’actualité [Lih, 2004c].
3. Vers un savoir co-construit, démocratique et perfectible ?

Le projet Wikipedia trouve tout son sens dans cette abolition potentielle de la frontière entre lecteur et auteur. Selon Arnaud Klein [Klein, 2005], la co-construction des connaissances repose sur une « collaboration à large échelle » avec un noyau dur de participants qui investissent spontanément des responsabilités éditoriales pour veiller à la structuration cohérente des contenus et garantir notamment la neutralité de point de vue (ou neutral point of view) et le respect du copyright. La constitution de la valeur s’appuie sur le partage des connaissances entre des participants motivés et capables de faire valoir leurs arguments pour accéder à une forme de neutralité et éviter les dérives sectorielles, politiques ou marchandes. Cette « vision dynamique d’un savoir en construction permanente » peut-il remplacer la vision circulaire d’une somme de connaissances universelles et absolues ?

Le modèle alternatif de Wikipedia est basé sur la régulation post-rédactionnelle et se construit au profit d’un contrôle partagé entre tous les participants. L’expression de tous les points de vue peut-être permettre d’atteindre la neutralité recherchée ? L’objectivité peut-elle se concevoir comme une somme de différences, voire de divergences sur certains sujets polémiques ? Autrement dit, « le partage, la coopération, l’intelligence collective sont-ils de nouveaux paradigmes qui permettent de construire les connaissances dans l’ère de la société de l’information ? Ou bien, s’agit-il d’une illusion, d’un phantasme relativiste ou utopique qui ne résistera pas, dans la durée, à la pression des faits et à l’exigence scientifique ? » [Klein, 2005].

Avec Wikipedia, le contrôle éditorial n’est pas seulement partagé, il est aussi démocratique et populaire. Il n’est pas doute le point névralgique à l’origine des réactions les plus violentes à l’égard de Wikipedia. C’est sans doute aussi ce qui remet le plus fondamentalement en cause le modèle traditionnel de validation par les pairs (peer review) : le consensus peut-il s’établir objectivement sur une majorité de contributeurs plus ou moins anonymes, plus ou moins sincères, voire plus ou moins uniques (cf. présence de sockpuppets : participants qui utilisent plusieurs comptes), votant pour telle ou telle vérité scientifiquement non certifiée ? Avec le groupe Intelligence Collective de la FING [Klein et al., 2006], on peut se demander si la moyenne des points de vue d’une masse importante de personnes, quel que soit leurs compétences, représente une information plus pertinente que le point de vue du meilleur des experts, comme le montrent par exemple les travaux de James Surowiecki [Surowiecki, 2005].

Les discussions initiées par Jean-Jacques Milan et Claire Panjel [GREMI, 2005] suite au signalement de Wikipedia dans le guide CERISE sont symptomatiques des questionnements sur la fiabilité de Wikipedia : à la question de savoir si elle peut être considérée comme une véritable encyclopédie et à ce titre être citée par les étudiants dans leurs travaux universitaires, la réponse des prescripteurs traditionnels est généralement négative.

Lors du séminaire GREMI de mars dernier [GREMI, 2005], Annaïg Mahé explique que le modèle de Wikipedia est d’autant plus questionnable que les auteurs ne sont pas identifiés (anonymat ou pseudonymie) et que le fonctionnement interne, notamment le processus de validation démocratique, reste flou et peu transparent pour l’internaute lambda.

Comme le souligne Larry Sanger [Sanger, 2004]*, Wikipedia souffre d’un déficit d’image. Sa popularité auprès des internautes ne change rien au fait qu’elle reste perçue comme non crédible par une majorité de bibliothécaires, d’enseignants et de chercheurs. L’absence de modération par ailleurs encourage des comportements polémiques stigmatisants chez certains contributeurs. Le crédit accordé par défaut à tout contributeur indépendamment de son expertise représente un préjudice majeur pour la légitimité de l’encyclopédie. La racine du problème est à chercher selon lui dans l’anti-élitisme et le manque de respect pour l’expertise inhérents au projet, voire à l’outil utilisé.

Beaucoup l’ont dénoncé par ailleurs : la présence de pertubateurs (trolls) et autres vandales forcent en quelque sorte les contributeurs experts à assumer des tâches défensives pour protéger leurs textes. « Wikipedia lets the vandals run loose on the creators, while commentators fill the void with chaff. It is a failure » [Scott, 2004]*. Et de tels affrontements seraient à l’origine de démissions fréquentes de la part des contributeurs sincères [Sanger 2004, 2005a et b]*. Une liste de Missing wikipedia est par ailleurs maintenues dans la Wikipedia anglaise.

Avec Wikipedia, l’auteur ne travaille plus dans l’isolement et ne peut plus s’appuyer sur sa notoriété pour légitimer son savoir, de même qu’il ne peut plus obtenir de reconnaissance immédiate avec la publication de son article. L’auteur est collectif et chaque individualité au sein de la communauté se doit effectivement d’argumenter pour défendre sa position et convaincre. Comment s’opère effectivement la coopération ? L’organisation de l’activité éditoriale relie-t-elle véritablement de l’anarchie, avec des contributeurs de plus en plus nombreux et de plus en plus réticents à toute remise en cause de cette culture du libre irrémédiablement ancrée comme le pense Larry Sanger [Sanger 2005a]* ? Peut-on identifier des fonctionnements bureaucratiques ? Qu’en est-il vraiment de l’exercice de la démocratie ? Peut-on parler de méritocratie pour qualifier la hiérarchie des pouvoirs ?

3.1. Quelques dispositifs pour l’auto-gestion et l’auto-régulation

Des principes fondateurs qui ont été arrivés à maturité dans les premiers mois de Wikipedia découlent différents dispositifs techniques et organisationnels mis en œuvre au sein des différents projets qui vont nous permettre d’apporter des éléments de réponse aux questions précédentes.

3.1.1. Dispositifs transversaux pour la gouvernance globale

La rapide croissance des encyclopédies régionales et l’émergence des projets jumeaux tels Wiktionary et Wikibooks ont nécessité la mise en place de dispositifs organisationnels et techniques pour structurer la gouvernance globale du projet.

Dispositifs organisationnels

Sur le plan organisationnel, cette gouvernance aujourd’hui est exercée par la Fondation Wikimedia, relayée dans les pays par des associations régionales.

- La Fondation Wikimedia

La création de la fondation Wikimedia en juin 2004 répond à un besoin d’institutionnaliser et de coordonner la gestion des projets basés sur le moteur MediaWiki à la fois sur les plans juridique, financier, managérial et technique. Les membres du comité sont élus via une consultation internationale à laquelle peut participer tout utilisateur inscrit depuis plus de 6 mois.

- Des associations régionales

Plusieurs associations ont vu le jour en 2005 pour soutenir le développement des projets Wikimedia : recherches de partenariats, mise en place d’un système de collecte de dons, communication externe : Wikimedia France, Wikimedia Deutschland et Wikimedia Italia.
L’édition de référence libre et collaborative : le cas de Wikipedia

Ces organisations offrent un début de réponse à la question du financement en proposant des structures habilitées à *recueillir des fonds* (donations de particuliers et d’organisations telles la fondation Lounsberry par exemple) et à négocier des partenariats pour l’hébergement des serveurs dans le monde entier.

En effet, les capacités de stockage des premiers serveurs payés par la société Bomis de Jimmy Wales n’ont pas résisté à la croissance exponentielle du trafic sur Wikipedia ; l’achat de nouvelles machines et la mise en place d’une infrastructure délocalisée sont devenues des priorités.

L’accord passé avec Yahoo pour l’hébergement gratuit des contenus de la Wikipedia anglaise en avril 2005, suite aux négociations engagées avec les deux géants américains Google et Yahoo, ne laissent pas d’interroger sur la neutralité d’un tel positionnement et la fragilité de ce modèle économique. La publicité fera-t-elle prochainement son apparition sur les pages de l’encyclopédie, alors que des sites commerciaux tels Answers redistribuent déjà les contenus de Wikipedia avec des annonces publicitaires ?

### Dispositifs techniques

- **Le Meta-Wiki** :

- **Des outils de communication extérieurs au wiki** :
  Selon Julien Levrel [Levrel, 2005], doctorant en sociologie des usages à l’université Marne La Vallée et membre du projet Autograph, les discussions concernant la gouvernance globale se sont progressivement déplacées vers des outils de communication asynchrone (*listes de diffusion*) ou synchrone (*canaux IRC* et messagerie instantanée) accessibles aux wikipédiens enregistrés. Certains canaux sont ouverts au grand public, d’autres sont réservés à quelques élus auxquels la collectivité a délégué un pouvoir de représentation (pour la négociation avec des partenaires extérieurs par exemple).

  Les canaux IRC constituent un « ciment social et communautaire » essentiel dans la contribution régulière à Wikipedia. Ce sont des espaces d’échanges, de négociation et de prise de décision, tant sur les contenus que sur les modes de gestion ; mais ce sont aussi des lieux de socialisation où les positions identitaires de chacun se confrontent. Le Village pump ou Bistro constituent quant à eux des lieux d’échanges informels publics au sein même de chaque wiki.

#### 3.1.2. Dispositifs techniques pour l’activité éditoriale

Wikipedia propose de nombreux outils utiles aux internautes pour suivre l’actualité de la communauté, prendre connaissance des modes de participation et assurer le suivi des articles. Leur caractéristique différentielle par rapport à d’autres outils d’édition réside dans leur visibilité et leur mise à disposition au sein même du wiki qui permet cette interdépendance entre contenus encyclopédiques et méta-informations. Nous en recensons ici quelques-uns uns parmi ceux les plus souvent mentionnés dans la littérature.

**Pages d’actualités, aides en ligne et outils de veille**

La page d’accueil de chaque version se présente comme un lieu hybride qui associe une présentation de l’encyclopédie, des espaces de médiation des contenus (actualités, article de la semaine, le saviez-vous, etc.), des outils d’accès aux contenus (recherche et navigation) et des informations sur la communauté wiki et l’administration du projet.

La page *Community portal* ou *Communauté* recense l’actualité de la communauté Wikipedia : en quelque sorte une sélection de pages pointant sur des tâches en attente de contributeurs, sur les prises de décision en cours au niveau régional et au niveau global, sur les projets en cours (accent sur le développement de portails thématiques ou la traduction d’un ensemble d’articles par exemple). Le moteur MediaWiki offre la possibilité à chaque contributeur de créer sa *page personnelle* (cf. *user pages*), afin de rendre publics son activité et ses domaines d’expertise.

Alors que le bac à sable ou sandbox permet aux contributeurs novices de faire leurs premiers pas sans risque, les rubriques d’aide fournissent un sommaire complet des conseils, conventions, modalités de participation et tutoriels disponibles (cf. Help contents ou Sommaire de l’aide) ; elles sont complétées dans la Wikipedia française par des pages *métasignales* signalant des ressources utiles au suivi des articles.

Le suivi des modifications peut s’opérer de différentes manières. A un premier niveau de coopération la page article donne accès à trois onglets (*discussion*, *modifier et historique*) assurant la traceabilité publique de la rédaction : le « code source » de l’article en quelque sorte. Elles se distinguent des pages articles par leur fonds jaune dans la Wikipedia française et ou bleu dans la Wikipedia anglaise.

Outre la possibilité avec les *watch lists* ou *listes de suivi* de s’abonner aux modifications d’une sélection d’articles en fonction de son activité éditoriale, chaque version propose des pages spéciales (cf. pages en anglais ou en français) utiles à la gestion des contenus : la page *Recent changes* ou Modifications récentes en particulier ou bien encore la page *Articles courts* dans la Wikipedia française. Les liens rouges signalent des articles qui n’existent pas encore, alors que les liens interlangues permettent d’accéder directement aux articles correspondants dans les autres versions linguistiques.

Il existe par ailleurs des agents automatisques ou semi-automatiques appelés bots qui interagissent avec Wikipedia comme le fait un utilisateur lambda mais pour des tâches répétitives et fastidieuses : gestion des liens interlangues, résolution des homonymies, opération sur les catégories, vérification de l’orthographe, identification d’actes de vandalismes, destruction des articles, etc. (cf. bots en français et en anglais).

**Modèles et labélisation des articles**

De nombreux modèles offrent du code prêt à être copié / collé pour créer de nouvelles pages. Un code couleur variable selon les versions régionales permet de distinguer les pages articles (fond blanc) des autres pages. Appliqués aux pages articles, ces modèles servent à la fois à standardiser la production des entrées de l’encyclopédie et à afficher des informations liées au cycle de vie de l’article.

Pour les articles les plus complets, la page s’appuie sur une structuration en chapitres et sous chapitres, sommaire, bibliographie, vignette illustrative et liens externes pour des compléments d’information. Certains portails thématiques proposent des modèles spécifiques pour normaliser la présentation et mieux gérer les renvois à l’aide de palettes de navigation (cf. par exemple, les modèles des portails *Géographie* ou *Domaine du vivant* dans la Wikipedia française et la liste complète des Navigational templates dans la version anglaise).
Ces modèles thématiques sont complétés par des modèles signatifs avec des étiquettes ou tags renseignant le visiteur sur le cycle de vie de l'article depuis ses conditions de création jusqu’à son statut actuel : Template messages ou Modèles pour la Wikipédia française.

Ces modèles servent à la promotion des articles avec notamment l’apposition d’un label Article de qualité ou Featured article qui permet d’identifier les articles promus dans cette catégorie par une icône spécifique placée en haut de page. Mais ils sont principalement utilisés pour présenter des Avertissements relatifs à la maintenance (tâches en cours ou à effectuer telles que problèmes de copyright, cas d’homonymie, titre incorrect, proposition de suppression, modification en cours, etc.), et à la gestion des controverses (désaccord de neutralité, suspicion de franco-centrage, conflit en cours d’arbitrage, etc.). Un projet d’harmonisation des messages d’avertissement est actuellement en cours dans la Wikipédia française.

3.2. Processus de stabilisation des articles et gestion des conflits

Les articles encyclopédiques évoluent dans des espaces où l’expertise rédactionnelle et l’expertise procédurale concourent dans un processus de stabilisation des contenus. Ce travail de veille ne correspond pas à un processus de validation des contenus en tant que tel, mais plutôt de réparation et d’évolution surveillée des contenus» [Levrel, 2005].

3.2.1. Coopération et gestion des conflits

Régulation, NPOV et WikiLove
Selon Andrew Lih [Lih, 2004c], l’efficacité de la coopération ne repose pas tant sur les caractéristiques techniques du wiki, assez pauvre en termes d’automatisation des tâches, que sur l’étalonnage des énergies créatives individuelles dans un espace où les interactions sociales sont valorisées. C’est cette « human orientation » qui favorise l’engagement personnel et l’investissement dans la communauté et qui permet à la régulation d’opérer via la recherche du consensus.


La neutralité de point de vue est mise au service de cet objectif à atteindre et fournit un cadre épistémologique pour la reconnaissance de tous les points de vue. La politique officielle de Wikipédia à l’égard des travaux scientifiques s’inscrit pleinement dans ce cadre initial (cf. No original research ou Travaux inédits) : l’encyclopédie n’est pas un lieu adéquat pour publier des travaux de recherche inédits.

L’autorégulation constitue un premier niveau de gestion des contenus suffisant pour permettre la co-construction d’articles. Cette autorégulation reste peu visible car elle fonctionne à un niveau local et s’appuie sur des espaces d’échanges externes au wiki lui-même. C’est quand l’autorégulation est mise en échec que le débat devient public : la régulation des conflits d’édition s’effectue par la consultation élargie de tiers participants. À un troisième niveau, la prise de décision repose sur des votes publics visant à exclure un ou plusieurs contributeurs. Dans la Wikipédia française, ce processus de d’arbitrage auquel la communauté a délégué le pouvoir de bannir les utilisateurs ayant gravement contrevenu à l’état d’esprit du WikiLove et/ou n’adhérant pas à l’objectif commun de l’horizon encyclopédique.

La création d’un nouvel article s’appuie généralement sur deux démarches inverses qui induisent des modes de collaboration différents, la deuxième restant fortement répandue dans les jeunes versions de Wikipédia :

- d’un côté un article d’emblée fortement structuré de la part d’un contributeur qui s’autoproclame expert : cet article est plus difficile à retravailler et nécessite une vérification du copyright
- d’autre côté une ébauche d’article ou un stub, qui entre immédiatement dans le circuit de valorisation, les quelques lignes introduitives servant à rendre public le besoin d’information : l’article se construit alors par petites touches

Le conflit d’édition (hors sujet extrême) est généralement bénéfique à la rédaction de l’article si chaque partie négocie en respectant l’état d’esprit du WikiLove. Lorsque le conflit se déplace du contenu pour toucher les personnes (agressivité, insultes), il se trouve toujours des contributeurs pour tenter une médiation et/ou rendre visible le débat en le transférant sur des pages publiques dédiées : Bistro ou Guerres d’édition en cours ou Conflits entre éditeurs par exemple. Ces procédures de publicisation peuvent conduire à une forte stigmatisation des individus à problèmes et aboutir très exceptionnellement à des sanctions. À cette dernière étape, ce sont les membres élus du Comité d’arbitrage qui interviennent pour prendre la décision finale. Pour connaître la politique officielle de la Wikipédia anglaise concernant la résolution de conflits, la page Resolving disputes peut être consultée.

Conflits et consensus
Au-delà de leur naïveté apparente, des principes tels que la neutralité de point de vue, le wikiplove et la wikipetiquette, jouent un rôle important dans l’acculturation des participants et dans la formation des modes opératoires, notamment pour ce qui concerne la résolution des conflits. Joseph Reagle [Reagle, 2005], diplômé du MIT en informatique et actuellement doctorant en écologie des médias à l’université de New York, examine par exemple la gestion des conflits dans Wikipédia à la lumière de concepts issus des recherches sur la prise de décision collective (interdépendant decision-making). Selon lui, les désaccords et conflits d’édition, bien qu’extrêmement minoritaires comparés aux actes de vandalisme, restent lourds de conséquences : leur résolution prend du temps et peut causer le retrait de certains participants ; ils influencent l’évolution des règles qui elles-mêmes affectent ensuite le reste de la communauté ; et pour finir leur publicisation attire exagérément l’attention à l’extérieur de la communauté.

Il remarque que le processus de résolution de conflits tel qu’il a pu être défini dans la littérature (cf. Carnevale et Pruitt notamment) est effectif dans Wikipédia. Ce processus fait appel à différentes stratégies issues de la négociation : la concession, la rivalité, la résolution de problème ou bien l’effacement. Si cette première étape de négociation informelle échoue, le débat nécessite une médiation (third-party decision-making), voire un arbitrage (joint decision-making).

Cependant le conflit, s’il peut être perçu comme un interdiction en termes d’interactions, doit être considéré dans son interdépendance avec la notion de consensus. Autrement dit, la coopération résulte d’une nécessaire fécondation entre accord et désaccord. La poursuite explicite d’un objectif commun dans Wikipédia comme produit de cette fécondation et est à rapprocher de la notion de communion utilisée par Howard Rheingold pour décrire les communautés virtuelles [Rheingold, 2005].

Civilité et politesse, recommandées par Wikipédia, font écho à la notion de perspective taking présente dans le management des connaissances. « By entertaining the perspectives of others, [participants] are able to engage in “perspective making”: identifying hidden commonalities and unique information in order to form a more valuable synthesis ». Ces pratiques sont également liées à la notion de dialogue basée sur l’empathie, popularisée par Daniel Yankelovich [Yankelovich, 2001].
3.2.2. Lutte contre le vandalisme

Dans une étude datée de 2003, Fernanda Viégas, Martin Wattenberg et Kushal Dave [Viégas et al., 2004a et b], de l'IBM Watson Research Center (Cambridge, Massachusetts), utilisent un outil de visualisation des flux de modifications (history flow) pour identifier 4 principaux scénarios collaboratifs : vandalisme et restauration / négociation / anonymat ou signature / stabilité.

Les résultats montrent que les actes de vandalisme se répartissent en 5 catégories : (1) mass deletion : suppression totale du contenu d'un article (2) offensive copy : insertion de mots vulgaires et insultes (3) phony copy : insertion d'un faux contenu sans relation avec le thème (4) phony redirection : insertion de fausses redirections (5) idiosyncratic copy : insertion d’un contenu contrefait, contreprenant à la NPOV. La comparaison des flux générés par des révisions successives espacées avec et sans la dimension temporelle a permis de constater que plus de la moitié des suppressions totales et des insertions de vulnérants sont corrigées en un temps inférieur à trois minutes.

La négociation se matérialise dans le flux par une série de zigzags qui se stabilisent après plusieurs échanges. Ces interactions correspondent à ce que Wikipedia appelle un conflit ou une guerre d’édition. Elles prennent fréquemment place dans les pages de discussion et ne sont pas limitées aux sujets traditionnellement controversés.

Les 70 entrées analysées par l’équipe de l’IBM Watson Research Center n’ont pas permis en revanche de relier la signature des articles aux sujets retenus d’une part, ni au volume des contributions d’autre part. Aucune relation claire n’a pu également être mise en évidence entre l’anonymat et le vandalisme.

La stabilité des contenus dans le temps reste aléatoire : les variations de volume et de texte sont continues. Cependant la première version d’une page tend à demeurer intacte plus longtemps, phénomène que les auteurs ont appelé « first mover advantage ». Une très large majorité des interventions correspondent à des ajouts ou des suppressions, et donc très exceptionnellement à des déplacements de texte dans l’article (un problème ergonomique ?).

Julien Levrel [Levrel, 2005] rappelle que plusieurs dispositifs permettent de détecter les actes de vandalisme dans la Wikipedia anglaise : la surveillance manuelle effectuée par les membres de la Recent Changes Patrol (RC Patrol) est complétée par une surveillance automatisée des modifications récentes à l’aide des Bots. Une fois détectés, les actes de vandalisme flagrants peuvent être corrigés directement par la restauration de la version précédente de l’article ; plusieurs tests ont montré que le temps de réaction est généralement inférieur à cinq minutes.

Pour des cas moins évidents ou graves, l’information doit être transmise aux administrateurs via le canal IRC dédié et faire l’objet d’une notification sur la page spéciale Vandalism in progress avec mention de l’adresse IP et du nom de l’utilisateur. Ils sont ensuite traités par les membres du Counter Vandalism Unit qui peut dans un premier temps placé un message d’avertissement sur la page suspectée et sur la page personnelle du contributeur quand elle existe.

La récente analyse de la page Vandalism in progress par Michael Lorenzen [Lorenzen, 2005], responsable des services de référence à l’université Central Michigan, lui permet de remarquer qu’un nombre très faible de comportements signalés ne relèvent pas à proprement parler du vandalisme tel qu’il est défini par Wikipedia, et sont alors attribués soit à une mauvaise manipulation, soit à un nouvel arrivant, soit à un signalement malveillant. Elle comporte une section dédiée aux sockpuppets, ces utilisateurs qui créent des identités multiples. Cette page permet de lancer un processus de surveillance qui peut aboutir à des mesures d’exclusion temporaire (de 4 à 48 heures dans le panel considéré) mais n’est pas représentative de l’ensemble des exclusions prononcées. Ce qui conforte en négatif d’autres résultats sur la multiplicité des méthodes mises en œuvre. [Ciffolilli, 2003 ; Levrel, 2005].

Pour en savoir plus sur les actes de vandalisme et la manière dont ils sont gérés, les pages Vandalism et Vandalisme peuvent être consultées.

3.3. Fiabilité et qualité

3.3.1. Tests et mesures quantitatives


W. Crawford pour sa part s’interroge sur la valeur de l’autorité en arguant que la validation par les pairs détermine finalement davantage la pertinence des recherches qu’il s’agit de publier plutôt que s’il va être publié ou non : « alternative publishing does not imply lack of worth. Traditional publishing isn’t an automatic indication of worth and veracity. If the key is “trust but verify”, we need better ways to verify likely worth and probity. »

L’actualité de ces dernières années ne manque d’ailleurs pas d’exemples, s’agissant d’erreurs commises dans des ouvrages de référence insouciables a priori : le Quid pointé pour ses articles révisionnistes ; Encarta et ses perles historiques ; les erreurs relevées dans le Dictionary of National Biography et celles de l’encyclopédie Britannica corrigées dans Wikipedia par exemple.

Andrew Lih [Lih, 2004a] mesure la qualité en termes de nombre d’éditions et de nombre de contributeurs uniques pour chaque article. En comparant une sélection d’articles avant et après leur citation dans la presse sur une période de 15 mois, il trouve que la promotion par le biais des citations dans la presse augmente de manière significative la « qualité » des productions. Il remarque par ailleurs que la majorité des articles cités sont fortement liés à l’actualité ou bien concernent des termes familiers ou argotiques.

Selon J. Levrel [Levrel, 2005], la progression parallèle du poids des articles (> à 2Ko) et de la création de nouveaux articles dans les Wikipedia anglaise et française indique clairement qu’une démarche qualitative est en cours, tendance qui pourrait s’expliquer par la meilleure visibilité du projet global et corollairement par l’arrivée de nouveaux participants. Si le taux de croissance des gros contributeurs (> 100 rédactions par mois) dans la version française reste plus faible que celui des participants, il permet d’identifier l’existence d’un noyau communautaire stable tout en pointant sur l’insuffisance des effectifs sur des tâches récurrentes d’administration et de gestion. L’examen de l’activité des 50 contributeurs les plus actifs (français et anglais) montre un engagement bipolaire différencié entre la contribution aux contenus et le contexte de justification des articles :

« Les différences de participation affichent des engagements participatifs complexes, où chaque participant construit son identité tant dans la construction d’articles que dans les débats internes qui fondent l’esprit de groupe et l’animation du projet, mais également des stratégies individuelles et de choix dans sa manière de s’engager ».

En outre, le nombre d’interventions nettement plus élevé dans la version anglaise permet d’établir un lien entre la taille du projet et l’émergence de phénomènes d’addiction pouvant correspondre à un engagement à plein temps de certains contributeurs.
3.3.1. Mesures qualitatives et construction temporelle

À ces études qui s’attachent à mesurer un mouvement quantitatif, il faut ajouter celle de Besiki Stvilia, Michael Twidle, Les Gasser et Linda Smith [Stvilia et al., 2005], tous quatre chercheurs à la Graduate School of Library and Information Science (University of Illinois, Urbana – Champaign), qui ont développé une approche à la fois quantitative et qualitative pour examiner le contexte social de production et la variation du niveau de qualité à l’aide d’une grille critériée très complète. En s’appuyant sur les différents rôles assumés par les participants, sur les pages de discussion et sur les articles de qualité (« featured articles »), ils identifient 10 types de problèmes liés à la qualité : information quality (scores, vérification), authority (manque d’expertise), completeness (taux de remplissage incomplet), complexity (manque de précision), inaccuracy (erreur de faits), consistency (manque de cohérence), informativeness (manque de consistance temporelle), relevance (manque de pertinence), verifiability (manque de références) et volatility (manque de stabilité).

Les problèmes les plus fréquemment pointés, tant pour les articles de qualité que pour le corpus aléatoire extrait de l’encyclopédie, sont accuracy et completeness, et dans une moindre mesure relevance, verifiability et consistency. À l’autre extrême, authority et volatility obtiennent les scores les plus bas. Selon eux, la majorité des problèmes de décontextualisation et apparaissent quand de l’information en provenance d’une source externe ne correspond pas au genre encyclopédique ou au contexte culturel et cognitif du lecteur lambda. Leur analyse souligne l’importance de ces problèmes de qualité pour les participants, conscients toutefois de l’existence de compromis dans leur traitement.

L’analyse de 120 articles proposés comme candidats au retrait du label « articles de qualité » (FARC pour Feature article removal candidates ) entre juillet 2004 et mai 2005 apporte un éclairage supplémentaire sur la construction temporelle de la qualité. Tout d’abord, les critères les plus fréquents pour labelliser un article et pour lui retirer son label sont identiques (completeness, verifiability, complexity and consistency) avec un score beaucoup plus élevé pour completeness et verifiability s’agissant des FARC. Ce résultat doit cependant être pondéré : la définition des standards de qualité répond à une maturation progressive de l’activité éditoriale. Ainsi le critère de verifiability a été ajouté en septembre 2004 et a été appliqué rétrospectivement aux articles labellisés. En outre, les auteurs notent que la communauté utilise le FARC pour redynamiser le processus éditorial quand des contributeurs estiment pouvoir améliorer le contenu d’un article.

En conclusion, les auteurs observent que la présence d’un petit groupe de contributeurs relativement homogène dans leur approche de la coopération est un facteur distinctif notable pour les « featured articles » et que le wiki valorise la construction du consensus en offrant des mécanismes d’interdépendance. Les pages de discussion contribuent à établir un sens de la communauté et à instaurer un ordre social basé sur le respect des normes de communication et en aidant les nouveaux participants à s’acculturer. Elles offrent par ailleurs aux lecteurs une possibilité de bénéficier de compléments d’informations sur certaines notions pas toujours accessibles aux néophytes, en consultant les discussions, voire même en posant directement des questions. Leur analyse montre en définitive que les participants assument cette approche constructiviste d’un savoir multiple et en constante évolution.

3.4. Dialectique wiki / contenus

Wikipedia est-elle simplement une encyclopédie basée sur un moteur wiki ou bien un wiki avec des contenus encyclopédiques ? Dans quelle mesure l’outil conditionne-t-il les modèles de collaboration et la nature des contenus ? Ouverture et décentralisation sont-elles des caractéristiques rédhibitoires à la production du savoir ? Comment cette disruptive technology concourt-elle à la création de la valeur ?

Alors que Ward Cunningham, initiateur du premier wiki en 1995, (cf. page dédiée à Wikipedia sur son site) considère que c’est l’outil qui dessine l’identité du projet, Larry Sanger estimait dans les premiers temps que le wiki n’était qu’un support pour organiser la collaboration. Il tend à croire aujourd’hui que les fonctionnalités d’un moteur wiki conditionnent fortement la manière dont les participants vont collaborer et consécutivement ce qu’ils vont produire ; si la culture wiki encourage une ouverture et une décentralisation totale, elle découvre également l’exercice de l’autorité et le respect pour l’expertise. Un outil réintroduisant une hiérarchie de rôles dans le processus éditorial permettrait de mieux contrôler la qualité des contenus et d’optimiser l’interdépendance. En s’appuyant sur les différents rôles assumés par les participants, sur les pages de discussion et sur les articles de qualité (« featured articles »), ils identifient 10 types de problèmes liés à la qualité : information quality (scores, vérification), authority (manque d’expertise), completeness (taux de remplissage incomplet), complexity (manque de précision), inaccuracy (erreur de faits), consistency (manque de cohérence), informativeness (manque de consistance temporelle), relevance (manque de pertinence), verifiability (manque de références) et volatility (manque de stabilité).

Les problèmes les plus fréquemment pointés, tant pour les articles de qualité que pour le corpus aléatoire extrait de l’encyclopédie, sont accuracy et completeness, et dans une moindre mesure relevance, verifiability et consistency. À l’autre extrême, authority et volatility obtiennent les scores les plus bas. Selon eux, la majorité des problèmes de décontextualisation et apparaissent quand de l’information en provenance d’une source externe ne correspond pas au genre encyclopédique ou au contexte culturel et cognitif du lecteur lambda. Leur analyse souligne l’importance de ces problèmes de qualité pour les participants, conscients toutefois de l’existence de compromis dans leur traitement.

L’analyse de 120 articles proposés comme candidats au retrait du label « articles de qualité » (FARC pour Feature article removal candidates ) entre juillet 2004 et mai 2005 apporte un éclairage supplémentaire sur la construction temporelle de la qualité. Tout d’abord, les critères les plus fréquents pour labelliser un article et pour lui retirer son label sont identiques (completeness, verifiability, complexity and consistency) avec un score beaucoup plus élevé pour completeness et verifiability s’agissant des FARC. Ce résultat doit cependant être pondéré : la définition des standards de qualité répond à une maturation progressive de l’activité éditoriale. Ainsi le critère de verifiability a été ajouté en septembre 2004 et a été appliqué rétrospectivement aux articles labellisés. En outre, les auteurs notent que la communauté utilise le FARC pour redynamiser le processus éditorial quand des contributeurs estiment pouvoir améliorer le contenu d’un article.

En conclusion, les auteurs observent que la présence d’un petit groupe de contributeurs relativement homogène dans leur approche de la coopération est un facteur distinctif notable pour les « featured articles » et que le wiki valorise la construction du consensus en offrant des mécanismes d’interdépendance. Les pages de discussion contribuent à établir un sens de la communauté et à instaurer un ordre social basé sur le respect des normes de communication et en aidant les nouveaux participants à s’acculturer. Elles offrent par ailleurs aux lecteurs une possibilité de bénéficier de compléments d’informations sur certaines notions pas toujours accessibles aux néophytes, en consultant les discussions, voire même en posant directement des questions. Leur analyse montre en définitive que les participants assument cette approche constructiviste d’un savoir multiple et en constante évolution.

3.4.1. Le Bazaar style : une approche décentralisée efficace

J. Levrel [Levrel, 2005] établit une filiation entre le développement des logiciels libres et les projets de contenus libres, en identifiant, outre les convergences idéologiques évidentes, plusieurs emprunts structurants :

- modèles d’organisation : agrégation de contributions individuelles bénévoles et hétérogènes (participation de non-experts notamment pour la remontée des bugs) au sein d’un tout et traçabilité de l’évolution,
- mécanismes de régulation : organisation distribuée associant expertise technique, expertise gestionnaire et expertise panoptique (surveillance globale et mise en cohérence),
- cadre juridique garantissant une ouverture du code (GPL : general public licence) et des manuels accompagnant les logiciels (FreeBSD, Documentation licence, Gnu Free Documentation Licence, etc.) en faveur d’une réusabilité.

Christian Wagner [Wagner, 2006], du département Information Systems à la City University of Hong Kong, va plus loin dans la parodie des approches qui favorisent les contributions multiples, plutôt que l’expertise de quelques-uns. Les résultats de l’étude menée par C. Wagner montrent que Wikipedia applique les méthodes et techniques du Bazaar style dans la création de connaissances en s’appuyant sur des modes conversationnels.
L’édition de référence libre et collaborative : le cas de Wikipedia

En effet, les articles de Wikipedia proviennent d’un développement par incrémentation et leur croissance et mode d’édition ne suivent pas la règle 80 / 20 de Vilfredo Pareto qui, appliquée à la gestion de projet, montre que 80% de la réalisation d’une tâche nécessite 20% de l’effort : 20% de l’existence moyenne d’un article comptent pour 35% de sa taille et 6% de son nombre d’éditions en 2001 (contre 15% en 2002, s’agissant du nombre d’éditions). La création et la maintenance des articles relèvent d’un effort collectif : de 18 à 285 contributeurs pour le panel constitué, avec une moyenne de 96 par article. La croissance de Wikipedia en termes de nombre d’articles est linéaire, voire super-linéaire, alors que les efforts dédiés à améliorer la qualité des articles augmentent avec le temps comme le montre la diminution du nombre de mots ajoutés à chaque nouvelle édition : 60 mots supplémentaires en moyenne à chaque édition pendant les 20% de son existence, contre 11 pendant les 80% restants.

3.4.2. Les systèmes à réputation remis en question

Des dispositifs de production collaborative tels que Slashdot ou Everything2 utilisent des systèmes de filtrage à la fois des contenus et des utilisateurs basés sur la popularité (reputation systems), un peu comme Amazon qui vous propose des ouvrages similaires en fonction des achats des internautes. Autrement dit, les articles les plus plébiscités par les votes de la communauté sont mis en avant et les auteurs de ces mêmes articles se voient gratifiés de plus de pouvoirs techniques pour continuer à rédiger des textes de qualité. Autrement dit encore, l’attribution par les utilisateurs d’une mauvaise réputation à un article affecte directement les droits de son auteur dans l’application. Ces systèmes à réputation se distinguent donc de Wikipedia sur ces deux points : les articles sont « signés » par un seul auteur et les droits à publier sont techniquement dépendants des votes.

Ces filtres collaboratifs postulent que la somme des expériences individuelles capitalisées bénéficie au renforcement « incrémental » de la qualité et de la cohérence des contenus produits par un contributeur unique : plus de crédits pour les « bons » contributeurs qui amendent leurs productions pour les rendre meilleures en fonction des votes de la communauté [Masum, 2004].

L’analyse comparée du niveau de standardisation et de structuration des articles dans les systèmes d’écriture collaborative utilisés par Wikipedia et Everything2, réalisée par William Emigh et Susan Herring [Emigh et al., 2005] (School of Library and Information Science, Indiana University), montre que plus le contrôle post-éditorial est ouvertement collaboratif et plus la ligne éditoriale est explicite, plus les articles présentent un degré de formalisation linguistique et structurelle fort.

Si les articles de Wikipedia ne se distinguent guère de ceux de la Columbia encyclopedia en termes linguistiques, ceux d’Everything2 font preuve d’un niveau de langue plus familier, voire humoristique. Paradoxalement donc, les entrées de Wikipedia sont stylistiquement plus homogènes et mieux organisées que celles d’Everything2 où le modèle éditorial est plus informel et où la régulation post-éditoriale repose sur le seul vote des participants, les contenus n’étant éditables que par leur auteur. S’agissant de Wikipedia, l’ouverture du système de collaboration et les pages de discussions semblent influencer positivement la production des entrées dans le respect de la neutralité de point de vue (NPOV) édicté comme principe structurant.
4. Les Wikipédiens : statuts, rôles et profils

4.1. Une hiérarchie de pouvoirs techniques

Quel que soit leur statut, tous les wikipédiens sont égaux en ce qui concerne l’activité d’édition (création et modification d’un article, participation aux discussions). Cette égalité de principe s’opère d’une manière collégiale entre tous les contributeurs, à travers une hiérarchie de pouvoirs techniques. Les administrateurs (« sysop ») ont le plus de pouvoirs, suivant de près les utilisateurs enregistrés. En revanche, l’utilisateur anonyme n’a que très peu de pouvoirs. Bien que la répartition des pouvoirs se fasse de manière collégiale entre tous les contributeurs, certaines tâches requièrent une forte acculturation technique et sociale au sein de l’environnement wiki et sont distribuées et donc restreintes à certains utilisateurs. Seuls les administrateurs – wikipédiens dont la candidature au poste d’administrateur a été validée par le collectif des contributeurs – sont autorisés par exemple à protéger, supprimer ou restaurer des pages et à bloquer des adresses IP ou des utilisateurs enregistrés. Ils sont plus de 70 dans la version française et près de 770 dans son homologue anglaise à l’heure de la rédaction (janvier 2006).

Le corps « professionnalisé » des administrateurs « regroupe des contributeurs ayant chacun des manières de s’engager très variées et un rapport au projet fortement individualisé. (…) Conservant un statut multiplexe qui lui permet d’intervenir à différents niveaux du projet en tant que participant, l’administrateur tend à structurer sa contribution en la spécialisant (page à déplacer / tags des photos / …) et en s’affichant comme porteur de la tâche (page personnelle, discussions wiki et IRC, …). D’une part, faire connaître que l’on se charge d’un travail peut enrichir l’identité du participant en légitimant sa carrière au sein du projet, et, d’autre part cela permet de canaliser les initiatives individuelles au sein d’un sous-projet si d’autres souhaitent également participer à ce travail » [Levrel, 2005].

Ces différents statuts peuvent être croisés avec des rôles plus spécialisés selon les tâches répétitives d’administration qu’ils choisissent d’exercer : bureaucrates, stewards, checkusers (vérificateurs d’adresses IP), médiateurs, arbitres. Ces rôles demeurent plus ou moins formalisés selon les versions et chaque participant peut bien évidemment investir plusieurs rôles. Le cas de Daniel Brandt (pseudonyme) par exemple, qui s’est illustré dernièrement en retrouvant l’auteur de la fausse biographie de Seigenthaler, montre d’une certaine façon que la traque des vandales et la publicisation des failles de sécurité est une spécialité comme une autre…

Pour aller plus loin sur ces questions, nous recommandons les pages Status et Functionaries pour la Wikipedia anglaise et les pages Statuts des utilisateurs, Différents rôles et Répartition des wikipédiens (par domaines de compétence, par année de naissance, mais aussi par heure de lever, etc.) pour la Wikipedia française. Les listes des contributeurs (avec renvoi sur leurs pages personnelles) sont consultables par rôle sur les pages suivantes : User list et Liste des participants.

4.2. Des architectes aux profils diversifiés

Selon Julien Levrel [Levrel, 2005], la gouvernance du projet Wikipedia « n’est ni basée uniquement sur l’expertise technique, ni sur l’expertise académique (rédaction de contenus), mais sur des carrières différenciées qui trouvent des arrangements quotidiens avec leur environnement immédiat ». Les éléments biographiques recueillis par questionnaires et entretiens auprès de contributeurs réguliers à la Wikipedia française lui permettent de dresser le portrait robot suivant :

- ce sont majoritairement des hommes, âgés de 25 à 49 ans, avec un niveau d’études moyen qui oscille entre bac +2 et bac +4 ;
- tous ont des expériences et pratiques très diverses de l’informatique, mais l’utilisent au quotidien soit dans leur vie professionnelle, soit dans leur vie privée, soit l’une et l’autre ; tous disposent d’un ordinateur personnel et d’une connexion Internet ;
- les sympathisants à la cause du logiciel libre sont par ailleurs fortement représentés, notamment les programmeurs ;
- étudiants et professionnels adoptent généralement une démarche pragmatique visant à réinvestir dans le projet des connaissances et des compétences qu’ils ont acquises ou qu’ils mettent en œuvre ;
- la majorité des Wikipédiens sont localisés sur le territoire français mais les rencontres physiques restent très exceptionnelles (signature des statuts pour l’association française par exemple) ;
- une forte majorité de contributeurs réguliers ont une expérience communautaire antérieure malgré des parcours hétérogènes (logiciel libre, jeux en ligne, newsgroups, forums de discussion, chat, etc.); cette pratique tend à se recentrer sur Wikipedia au détriment des autres communautés ;
- la captation de l’individu repose prioritairement sur des éléments biographiques personnels : la nature du projet et les types d’échanges répondent à des attentes individuelles ;
- corollairement la phase d’apprentissage et la routinisation des procédures de coopération sont plus importantes que l’expertise académique pour légitimer sa carrière au sein du collectif ;
4.3. Des écoles de pensée qui s'affrontent

Ce panorama des profils d'utilisateurs resterait incomplet si nous ne mentionnions pas l'existence quasi corporatiste des Metapedians, sorte d'équipes plus ou moins organisées et plus ou moins controversées, représentant un vaste panel de philosophies liées au projet même et à son évolution. Les écoles de pensée dominantes sont structurées en association et disposent de leur propre page affichant la liste de leurs membres, leur charte et leurs principes sur un ton souvent humoristique, mais qui ne manque pas de révéler toute l'importance que revêt Wikipedia à leurs yeux. À une extrémité du spectre philosphique, on trouve l'Association of deletionist wikipedians qui rêve d'une encyclopédie expurgée de tout contenu malpropre et qui favorise l'établissement de standards rigoureux pour accepter un article, voire le développement d'une wikilaw, sorte de jurisprudence. A une autre extrémité, l'Association of inclusionist wikipedians privilégie l'amendement des articles problématiques plutôt que leur suppression et mise sur l'activité organique de la communauté ; ils constituent la branche liberatoire des Metapedians et placent la lutte contre les rampants deletionists au cœur de leur activité. Au centre, l'Association of mergist wikipedians s'investit d'une mission de réconciliation et prône le juste milieu entre deletionists et inclusionists. S'il est difficile aujourd'hui de mesurer l'influence de ces bandes, leur existence reste révélatrice de la dynamique interne au projet et en quelque sorte de la pluralité identitaire qui le compose [Vershbow, 2006].

4.4. Acculturation et création de sens

La croissance exponentielle des participants dans une communauté virtuelle peut menacer sa capacité à évoluer et rendre la gouvernance impossible, comme l’ont montré certaines études sur les forums de discussion. La manière dont les nouveaux arrivants s'intègrent dans le projet représente donc un enjeu essentiel pour son évolution.

Cormac Lawler [Lawler, 2005c], auteur d'un mémoire de master intitulé *Wikipedia as a learning community* à l’université de Manchester (School of Education in Communication, Education and Technology), rappelle que si les qualificatifs de monarchie, d'anarchie, de démocratie ou bien encore de bureaucratie font débat, celui de méritocratie s'avère généralement plus consensuel s'agissant de Wikipedia. Cependant l'émphase dans le modèle méritocratique doit être tempérée par une reconnaissance de l'attitude à encourager la pratique réflexive. La qualité du discours prend dès lors une importance cruciale, supérieure à celle des contributions mêmes, pour encourager cet « educational process ». L'acculturation des nouveaux arrivants implique en effet une réflexion critique de chacun à la fois au modèle du praticien réflexif [Schön, 1990] selon le modèle du praticien réflexif [Schön, 1990].

Susan Bryant, Andrea Forte et Amy Bruckman [Bryant et al., 2005], chercheurs au département informatique du Georgia Institute of Technology, s'intéressent au processus d'acculturation des contributeurs en Wikipedia en appuyant leur analyse sur les perspectives offertes par la théorie de l'activité (activity theory ou AT) et par la participation périphérique légitimée (legitimate peripheral participation ou LPP), tout en se référant aux précédents travaux sur les communautés de pratique (communities of practice ou CoPs) conduits par Etienne Wenger et Jean Lave [Lave et al., 1991]. Les six éléments issus de la théorie de l'activité (objet, sujet, communauté, division du travail, outils, règles) fournissent donc un cadre méthodologique pour évaluer la LPP et l'évolution de la nature des participations et pour comprendre comment les participants créent du sens dans leurs pratiques.

Dans leurs premières contributions, les sujets se comportent comme des consommateurs qui opèrent occasionnellement et fortuitement quelques changements mineurs sur les articles qu'ils lisent, avant d'adopter un rôle de gardien au regard d'une collection d'articles correspondant à ce qu'ils connaissent. Puis leurs objectifs s'élargissent et ils deviennent progressivement plus concernés par l'ensemble de l'encyclopédie et par le développement de la communauté, plutôt que par la qualité de quelques articles isolés. Si les wikipédiens croient au projet de l'encyclopédie et au travail de la communauté, leurs motivations ne sont pas à chercher du côté de l'altruisme, comme plusieurs travaux sur les CoPs l'ont démontré, mais bien du côté d'un positionnement identitaire qui s'inscrit dans le temps.

L'évolution des pratiques touche également l'utilisation de l’outil. À un premier niveau, la facilité d'édition via la fonction « modifier la page » joue un rôle important dans la transition entre le lecteur-consommateur et le lecteur-éditeur. L’élargissement de leur engagement et donc de leur zone proxime de développement (zone of proximal development ou ZPD) rend certains outils plus visibles parce que plus porteurs de sens, telles les fonctions « historique » et « discussion » non utilisées par les novices. L'utilisation des listes de suivi se ritualise, pour s'étendre à la surveillance globale des modifications récentes et des nouvelles pages.

Au départ, Wikipedia est perçue comme une collection d'articles plutôt que comme une collection d'individus. Les novices n'ont pas conscience des différents rôles associés à la division des tâches, contrairement aux CoPs traditionnelles décrites par Etienne Wenger. Les seules règles dont ils ont connaissance sont liées à l'édition d'une page. Au contraire, les wikipédiens se considèrent comme membres d'une communauté et l'anonymat devient suspect. Les pages personnelles sont un lieu adapté à la fois pour fournir quelques éléments biographiques et présenter ses activités au sein du projet. Leur connaissance des règles se précise, allant des principes fondateurs de la communauté aux conventions syntaxiques et procédures techniques de gestion des pages. La reconnaissance du même identité s'exprime dans les pages de discussion attenantes aux articles, et notamment dans celles liées aux pages personnelles, alors que les rôles se diversifient et se spécialisent au fur et à mesure que l'investissement personnel se renforce.
Le focus de S. Bryant, A. Forte et A. Bruckman sur le processus d’acculturation apporte un éclairage complémentaire à celui de J. Levrel [Levrel, 2005] plus tourné vers les profils et motivations des participants. Ces deux études ouvrent des perspectives nouvelles même si l’échantillon des personnes interrogées est d’une part trop faible pour être représentatif et d’autre part focalisé sur des participants déjà fortement impliqués dans l’activité de la communauté. De même, l’enquête actuellement en cours à l’université de Wuerzburg en Allemagne (département psychologie) s’intéresse aux motivations des contributeurs actifs [Schroer et al., 2005]

Si la variable générationnelle est généralement peu distinctive dans la communication médiatisée par ordinateur (computer mediated communication), les identités sociales mériterait d’être questionnées plus en profondeur, notamment pour ce qui concerne l’influence des variables professionnelles et des contextes linguistico-culturels : identifier par exemple la proportion de chercheurs, bibliothécaires ou enseignants et investiguer leurs usages, mettre en relation les domaines de compétences et l’activité éditoriale... Les résultats de l’enquête réalisée récemment pour la revue Nature [Giles, 2005] montrent que moins de 10% des chercheurs interrogés disent avoir déjà contribué à Wikipedia : une proportion estimée tout à fait insuffisante par l’éditeur qui encourage explicitement ses lecteurs à s’engager dans cette aventure éditoriale au bénéfice d’un savoir partagé dynamique. Dans son éditorial Jim Giles suggère que l’échange entre expert et non-expert doit pouvoir s’avérer stimulant pour les deux parties, et en particulier pour l’expert qui relève le défi d’une autre forme de communication scientifique.

4.5. Contributions périphériques ou anonymes

De même, l’étude des freins à la participation et des causes de retrait apporterait sans doute des éclairages intéressants sur le fonctionnement organique de Wikipedia. Que se passe-t-il pour cette majorité silencieuse qui reste finalement à la périphérie du projet ? Et pour cette minorité impertinente et peu respectueuse des principes de base ? Citée par Cormac Lawler [Lawler, 2005b], la « wiki editing dichotomy » de Chris Allen, se vérifie-t-elle ? : « One interesting possible barrier of entry to active participation in a wiki is what I call the “wiki editing dichotomy”. You have to be proud enough to believe what you are contributing is generally worthwhile to others (or at least worth your effort), but you also have to be humble enough to understand that others can improve it. I don’t know of many other collaborative media that requires both pride and humility. »

En contrepoint, les modalités de participation des contributeurs enregistrés et des contributeurs anonymes restent insuffisamment explorées. Tout se passe comme si ces formes d’engagement n’étaient pas considérées comme discriminantes. Même si l’étude de l’IBM Watson Research Center [Viegas et al., 2004a et b], ne permet pas d’établir de relations entre anonymat et vandalisme, la proportion d’éditions anonymes qui s’élevait fin 2004 à 42% dans la Wikipedia japonaise, contre 10%, 15%, 22% et 26% respectivement dans les versions italienne, française, anglaise et allemande [Voss, 2005a] interroge immédiatement sur les styles de contribution à l’œuvre.

Les travaux de Denise Anthony, Sean Smith et Tim Williamson [Anthony et al., 2005] (Dartmouth College, Hannover), apportent cependant un début de réponse à ces questions sur l’anonymat. Orientée sur la qualité des contributions entre participants enregistrés et participants anonymes, leur analyse fait apparaître des styles d’engagement qualitatif diversifiés selon la proportion de contenus retenue après une intervention sur un article et selon le niveau de participation. A partir d’un échantillon de 7000 individus issus des Wikipedia allemande et française, ils montrent qu’à côté de ces contributeurs zélés, motivés par l’acquisition d’une bonne réputation et/ou par l’adhésion à un projet communautaire, il existe un grand nombre de « bons samaritains » anonymes qui produisent très occasionnellement et de manière ciblée des contenus de qualité. Dans cette configuration, le maillon faible est représenté par les contributeurs anonymes dont la fréquence des interventions est élevée :

« What are the implications for quality for anonymous contributors? Virtually all theories of social dilemmas would predict low quality contributions from anonymous contributors, especially those with low levels of participation, since they would seem to have little motivation or incentives to contribute. Yet the lore of open-source suggests that anonymous one-time contributors are as important as the zealots. Who are these Good Samaritan contributors? They are likely to be of two types. The first type of Good Samaritans may be, like the zealots, experts in a particular field. These experts do not care about their reputation in Wikipedia (no registration), nor are they committed to Wikipedia as a community (few contributions). Instead they care about their area of expertise and so contribute to that topic only. Taking the time to register would actually increase the costs of contributing for these Good Samaritan contributors, and since they are not interested in reputation and do not identify with the community itself, they have no reason to incur these costs. Given their expertise in the subject matter, however, their contributions will be of high quality. Alternatively, the second type of Good Samaritan contributors are likely to be merely passers-by who see a mistake or a hole and make a contribution to address it. These contributions are likely to be much shorter than others, and therefore less likely to be edited or changed in the future, making them appear high quality. Thus, in contrast to registered users whose quality is highest at high levels of participation, anonymous users with the fewest contributions will be the highest quality. »

Pour les auteurs, les contributions de ces « bons samaritains » ne peuvent être considérées comme un éphéméromène et constituent à part entière une donnée significative de nouveaux usages relatifs à la production collaborative de contenus libres, plaçant ainsi l’innovation non seulement à un niveau technique mais aussi à un niveau organisationnel. Ces résultats mettent donc en lumière de nouvelles modalités de production méconnues et donc peu exploitées, tout en relatissant le discours dominant sur les motivations liées à la participation à un projet communautaire et les théories traditionnelles sur le fonctionnement des CoPs. Ces résultats, en définitive, relancent le débat sur l’instauration d’une authentification préalable obligatoire pour la participation au wiki.
5. Vers une organisation apprenante ?

Si l'on admet la nécessité pour le lecteur d'adopter une attitude critique à l'égard des contenus de l'encyclopédie (et à l'égard de toutes les sources par ailleurs), il convient de reconnaître que la contribution repose aussi sur une « action critique participante » visant à rendre l'encyclopédie « perfectible par itérations successives ». [Klein, 2005]. Là où Arawnda Klein parle de « posture journalistique », certains évoquent « compétences informationnelles », « formation à l'information », « éducation aux médias », etc.

Si les non-experts en documentation tendent à se focaliser sur l'information brute en questionnant superficiellement les éléments contextuels et conditions de production et si la démarche d'évaluation de l'information en ligne requiert un apprentissage, comme l'ont montré les récents travaux de l'ERT IRMA et du LaCo [GREMI, 2005 : présentation de Monica Macedo Rouet], pourquoi ne pas miser sur cette « action critique participante » pour développer ces compétences ? Wikipedia serait-elle un environnement pédagogique adapté si l'on considère avec J. Levrel [Levrel, 2005] que : « En perpétuelle évolution et co-construction, la valeur de l'article ne s'acquiert, à l'œil du lecteur, qu'après une démarche personnelle, critique et réfléchie de mobilisation de traces de construction. » ? Autrement dit, des positions divergentes au départ trouvaient-elles sur le terrain de la participation à Wikipedia un lieu de réconciliation ? Andy Carvin l'a imaginé [Carvin, 2005*], d'autres l'ont fait.

5.1. Revue de projets pédagogiques


Avec le projet Égypte lancé à la rentrée 2005 dans l'externat St-Jean Berchmans (Québec), Mario Asselin fait le pari que la participation des élèves à un projet authentique d'écriture collective stimulera leurs motivations et donc leur apprentissage. Les contenus seront élaborés dans un premier temps sur le wiki propre à l'école primaire puis transférés sur Wikipedia quand ils auront atteint une maturité suffisante. Dans cette expérience, la publication sur Wikipedia ne constitue pas une fin mais plutôt un moyen pour encadrer l'apprentissage dans un contexte réel et déconstructeur. Dans une veine, nous retrouvons deux projets universitaires d’envergure. Dans celui de l’Indiana University animé par Kevin Zollman, des étudiants de philosophie ont reçu pour consigne de vérifier les contenus de quelques articles à l’aide de ressources papier avant de motiver leur intervention dans Wikipedia. Dans le cadre d’un cours sur les nouveaux médias à l’université de Hong Kong, Andrew Lih propose quasiment chaque semestre à ses étudiants d’éditer une collection d’articles sur un sujet précis : WTO (OMC) était le thème retenu dans le dernier atelier.

L’encyclopédie ne se contente pas de se nourrir de ces exercices. Selon la logique de publication réflexive qui lui est propre, elles les donnent aussi à voir… et à penser : School and university projects, Wikipedia in der schule, Projets pédagogiques. Et la communauté s’en empare en prodiguant conseils et recommandations aux enseignants.

Parallèlement à ces projets pédagogiques, des démarches ouvertes de promotion sont initiées au sein des instances éducatives. Ainsi, au Québec toujours, le RECI T (réseau pour le développement des compétences par l’intégration des technologies) a ouvert dernièrement un site qui médiatise les programmes scolaires à l’aide d’une liste de concepts renvoyant sur Wikipedia. Créé dans le cadre d’un projet de formation de l’école québécoise pour l’éducation primaire et secondaire, le site PDFEQ-MST-Wikipedia se présente donc comme une « baie ouverte » sur Wikipedia pour les mathématiques, les sciences et les technologies (Létourneau, 2005*).

D’autres projets au sein de la fondation Wikimedia intéressent ouvertement enseignants et apprenants. Les wikiversités ou wikiversités intégrées qui sont de nouveaux projets dewikis ou Wikilivres en français, sont des espaces dédiés à la publication de cours et de manuels ; elles sont structurées en facultés : School of education par exemple.

La société Wikicités fondée par Wikimedia offre par ailleurs l’hébergement de projets wiki à toute communauté intéressée ; le projet Education Wikiacyrège par exemple des wikis universitaires multilingues en cours de déploiement. Cette revue d’initiatives que l’on pourrait qualifier de « citoyennes » questionne bien évidemment notre rapport à la connaissance et soulève de nombreuses questions quant aux modèles d’apprentissage en œuvre. Dans ces expériences, le savoir ne se transmet plus : il se construit et même il se co-construit. Cette vision dynamique, mise explicitement au service de l’apprentissage, est-elle effective ?

5.2. Wikipedia comme organisation apprenante ?

5.2.1. Retours d’expérience sur deux projets pédagogiques

En 2004 à la Queensland University of Technology, 150 étudiants issus de 6 classes différentes ont réalisé une collection de textes encyclopédiques sur les nouveaux médias – MCyclopedia of new media – à l’aide du moteur MediaWiki. Les résultats de cette expérience [Bruns et al., 2005] sont contrastés. Si les étudiants se sont très rapidement appropriés l’environnement technique, ils ont cependant rencontré plus de difficulté avec le style encyclopédique : les premières contributions ressemblaient davantage à des travaux de rédaction type thèse – antithèse – synthèse, plutôt qu’à un ensemble de points de vue présentés dans une perspective de neutralité.

De même, la coordination entre les classes via le wiki s’est avérée peu opérationnelle alors que les échanges au sein d’une même classe ont dans l’ensemble bien fonctionné. Le manque de fonctionnalités relatives à la désambiguisation dans le moteur wiki ont par ailleurs accentué ces problèmes de coordination pour le nommage des articles.

Le nombre restreint de participants comparé à un projet tel que Wikipedia n’a pas permis une évaluation satisfaisante de la dynamique de co-construction tant au niveau de la correction des articles que de l’organisation des contenus. Dans une certaine mesure, l’homogénéité des profils des participants a également desservi la confrontation des points de vue en introduisant en quelque sorte des biais culturels dans le traitement des sujets. Le projet pourrait être élargi à d’autres classes issues d’autres universités pour palier – du moins partiellement – à ces difficultés.

Comme le projet Égypte évoqué précédemment, celui du Center for Teaching Learning, and Technology (CTLT) de la Washington State University est d’abord développé sur un wiki local pour permettre une meilleure appropriation de l’outil MediaWiki et pour favoriser une collaboration plus rapprochée avec l’équipe enseignante, difficilement envisageable dans l’encyclopédie. L’exercice est décomposé
5.2.2. Action critique et réflexivité explicite

Cormac Lawler [Lawler, 2005a] rejette le postulat de l’environnement sécurisé en positionnant le conflit d’édition comme une composante-clé à la fois du processus de création des contenus et de la participation des contributeurs. Selon lui, Wikipedia n’est pas simplement une encyclopédie, ni même un nouveau modèle éditorial : le projet présente en effet les caractéristiques d’une organisation apprenante tant dans sa structure que dans son fonctionnement. A la lumière des recherches sur les communautés virtuelles ou les communautés de pratique, il identifie les éléments de convergence suivants : Wikipedia est auto-gérée, auto-sélective et auto-évolutive. Le potentiel de la communauté réside selon lui dans la promotion d’une pédagogie de la découverte et dans le développement de l’esprit critique nécessaire à son fonctionnement. Le conflit est perçu comme un ingrédient nécessaire à la négociation car il requiert une mise en perspective des différents points de vue et une prise en compte des différences culturelles, comme le montre le cas de controverse anglo-franco-canadienne dont il fait état. La confrontation des partis-pris culturels (cultural biases) est un signe d’action critique manifeste dans un contexte où l’adhésion à la règle de neutralité est partagée et où la convivialité est assumée comme essentielle.

Dans ses observations [Lawler, 2005a], il a en effet constaté que les participants expriment le sentiment d’appartenir à une communauté et mentionnent cette convivialité (sense of welcome) et cet état d’esprit d’aide réciproque (mutual aid) consubstantiels au projet. Consensus et désaccord sont perçus comme vitaux pour la réalisation du produit final, bien que le consensus une fois atteint représente un frein à de nouvelles contributions.


Dans Wikipedia, cette pratique réflexive est incorporée et s’exerce à différents niveaux, comme nous l’avons vu : pages de discussion liées aux articles, pages métas sur les règles et conventions, pages métas sur les usages (cf. supra) et sur les performances (statistiques et tableaux de bord), pages bibliographiques sur les citations et les travaux de recherche…), etc. Selon C. Lawler, il est essentiel que les nouveaux arrivants bénéficient de cette expérience et donc que la mémoire collective continue à être explicitée (via des pages de type lessons learnt par exemple). De même, l’apprentissage par l’expérience et l’expertise cumulée et en quelque sorte collectivisée permet de développer des compétences qui seront essentielles dans le monde professionnel de demain.

Ces réflexions font largement écho à ce qu’il est convenu d’appeler maintenant le e-learning 2.0 dont Stephen Downes représente l’un des théoriciens les plus actifs :

« What happens when online learning ceases to be a type of content consumption tool, where learning is “delivered”, and becomes more like a content-authoring tool, where learning is created? The model of e-learning as being a type of content, produced by publishers, organized and structured into courses, and consumed by students, is turned on its head. Insofar as there is content, it is used rather than read—and is, in any case, more likely to be produced by students than courseware authors. And insofar as there is structure, it is more likely to resemble a language or a conversation rather than a book or a manual » [Downes, 2005].

Selon Jean-Claude Guédon, on passe alors d’un modèle de consommateur passif de connaissances à celui de citoyen acteur de sa propre connaissance [cité par Asselin, 2005]."
Conclusion

Evaluer Wikipedia en termes d’échec ou de réussite reste largement subjectif si, plutôt que de la considérer comme un produit éditorial concurrent des encyclopédies telles que Britannica et Universalis, on la perçoit comme une dynamique de projet autour d’un objectif visant un « horizon encyclopédique ».

Les débats sur la légitimité de Wikipedia ne sont d’ailleurs pas sans rappeler ceux qui ont marqué l’histoire de l’Encyclopédie de Diderot et d’Alembert, perçue comme partielle et extrêmement dérangeante à l’époque, tant pour ses contenus que pour ses modes de diffusion.

Pour autant, cette revue de littérature scientifique nous permet de souligner l’opérationnalité globale de l’auto-régulation et du processus de stabilisation des contenus ; toutes font état de cette tension vers la qualité inhérente à sa dynamique interne. Aucune, en revanche, ne permet de pointer une relation entre anonymat et vandalisme par exemple.


Mais au-delà des controverses, le soutien à cette démarche qualitative apparaît presque comme une évidence : c’est l’usage qui crée l’innovation et nul doute que cette « innovation ascendante » [Guillaud et al., 2005] continuera à remonter pour toucher experts et chercheurs qui y trouveront des opportunités nouvelles pour partager leur savoir et un terrain « grandeur nature » inédit pour l’étude des communautés virtuelles. Mais elle continuera parallèlement à s’étendre en attirant d’autres « non-experts » aux profils variés (et donc d’autres vandales et d’autres perturbateurs).

Il ne s’agit donc pas de faire preuve d’un optimisme béat. La réussite des innovations ascendantes dépend aussi de l’articulation dialectique entre « contenant » et « contenu », comme l’a montré Eric Von Hippel [Von Hippel, 2005]. La croissance, nous l’avons vu exponentielle, à la fois des contenus et des participants ne risque-t-elle pas d’atteindre un seuil critique où l’équilibre précaire entre les évolutions techniques du wiki et la production des contenus sera rompu ? Peut-on craindre que des forces d’inerties, voire des formes de résistances au changement deviennent suffisamment présentes pour figer ces deux dynamiques ?

La régulation post-éditoriale pourra-t-elle continuer à opérer et à progresser dans un environnement complètement ouvert et décentralisé alors que contributeurs et contributions vont manifestement continuer à augmenter ? Beaucoup s’interrogent déjà sur la manière dont la communauté va négocier ce tournant de son histoire et faire face à la nécessaire redéfinition de son projet initial, comme le montrent les réflexions actuelles sur l’édition de versions stabilisées. Généraliser les indicateurs de qualité sera-t-il suffisant pour procéder à cette redéfinition ? Faudra-t-il introduire une forme de validation qui reconnaisse l’expertise et bannisse l’anonymat ? Faudra-t-il privilégier la tolérance envers la diversité des formes d’engagements que peuvent investir les contributeurs ? Toutes ces questions et bien d’autres ne manquent pas de nous interpeller et d’offrir de nombreuses pistes à explorer…
Bibliographie

Toutes les références, enrichies de résumés et de liens sont consultables et téléchargeables sur notre bibliographie collaborative
Encyclo : http://wikindx.inrp.fr/biblio_encyclo/

Recherches


Giles, Jim. 15 décembre 2005. Internet encyclopaedias go head to head. News@nature.com.

Gillmor, Dan. 2006. We the media: grassroots journalism by the people, for the people. 1re éd. O'Reilly Media, Inc. 0-596-10227-5.


L’édition de référence libre et collaborative : le cas de Wikipedia


Débats et réflexions


L'édition de référence libre et collaborative : le cas de Wikipedia

Sitographie


Conférences, manifestations

Wikis et autres systèmes collaboratifs
- [http://c2.com/](http://c2.com/)
- [http://slashdot.org/](http://slashdot.org/)
- [http://susning.nu/](http://susning.nu/)

Encyclopédies libres ou projets concurrents
- [http://enciclopedia.us.es/](http://enciclopedia.us.es/)
- [http://planetmath.org/](http://planetmath.org/)
- [http://susning.nu/](http://susning.nu/)
- [http://uncyclopedia.org/wiki/Main_Page](http://uncyclopedia.org/wiki/Main_Page)
- [http://www.bbc.co.uk/dna/h2g2/](http://www.bbc.co.uk/dna/h2g2/)
- [http://www.disinfopedia.org/](http://www.disinfopedia.org/)
- [http://www.editionsatlas.fr/files/453510d0.pdf](http://www.editionsatlas.fr/files/453510d0.pdf)
- [http://www.jurispedia.org/](http://www.jurispedia.org/)

Erreurs relevées dans d’autres encyclopédies
- [http://en.wikipedia.org/wiki/Wikipedia:Errors_in_the_Encyclop%C3%A6dia_Britannica_that_have_been_corrected_in_Wikipedia](http://en.wikipedia.org/wiki/Wikipedia:Errors_in_the_Encyclop%C3%A6dia_Britannica_that_have_been_corrected_in_Wikipedia) (Britannica)
- [http://observer.guardian.co.uk/uk_news/story/0,6903,1431473,00.html](http://observer.guardian.co.uk/uk_news/story/0,6903,1431473,00.html) (Dictionary of National Biography)
- [http://www.cvm.qc.ca/APHCQ/dossiers/encarta.htm](http://www.cvm.qc.ca/APHCQ/dossiers/encarta.htm) (Encarta)
- [http://www.phdn.org/negation/quad/quad.html](http://www.phdn.org/negation/quad/quad.html) (Quid)

Modes d’accès et de consultation pour encyclopédies et dictionnaires
- [http://answers.google.com/answers/](http://answers.google.com/answers/)
- [http://autograph.fing.org/](http://autograph.fing.org/)
- [http://cellphedia.com/](http://cellphedia.com/)
- [http://conquery.mozdev.org/](http://conquery.mozdev.org/)
- [http://dict.mozdev.org/](http://dict.mozdev.org/)
- [http://dictionarysearch.mozdev.org/dictionaries.html](http://dictionarysearch.mozdev.org/dictionaries.html)
- [http://mobile.answers.com/](http://mobile.answers.com/)
- [http://mycroft.mozdev.org/quick/wikipedia.html](http://mycroft.mozdev.org/quick/wikipedia.html)
- [http://wiki.netbib.de/coma/Linky](http://wiki.netbib.de/coma/Linky)
L’édition de référence libre et collaborative : le cas de Wikipedia

- http://wordnet.princeton.edu/
- http://www.answers.com/
- http://www.dicoweb.levillage.org/
- http://www.encyclopedia.com/
- http://www.factbites.com/
- http://www.factmonster.com/
- http://www.hyperwords.net/
- http://www.i-km.com/alexandria.htm
- http://www.infoplease.com/
- http://www.onelook.com/
- http://www.placeopedia.com/
- http://www.quero.at/
- http://www.wapipedia.org/
- http://www.xrefer.com/
- http://yubnub.org/

Statistiques de fréquentation et d’usages de Wikipedia

- http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Communique%C3%A9_de_presse_pour_les_100_000
- http://meta.wikimedia.org/wiki/Wikipedia.org_is_more_popular_than...

Usages des wikis et de Wikipedia

- http://wikiesjb.opossum.ca/ProjetEgypte
- http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Projets%2C%3A9dagogiques
- http://www.recit.qc.ca/
- http://recitmst.qc.ca/pdfeq/
- http://fr.wikibooks.org/wiki/Wikiversity%C3%A9
- http://education.wikicities.com/wiki/Main_Page
L'édition de référence libre et collaborative : le cas de Wikipedia


**profils des Wikipédiens**

- http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Les_diff%C3%A9rents_r%C3%A9porteurs
- http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:R%C3%A9partition_des_wikip%C3%A9diens
- http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Statuts_des_utilisateurs
- http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Wikip%C3%A9diens_par_domaine_de_comp%C3%A9tence
- http://meta.wikimedia.org/wiki/Inclusionism
- http://meta.wikimedia.org/wiki/Mergism
- http://meta.wikimedia.org/wiki/Metapedianism
- http://meta.wikimedia.org/wiki/Status
- http://wikipediareview.com/

**gouvernance globale**

- http://commons.wikimedia.org/wiki/Accueil
- http://mail.wikimedia.org/
- http://meta.wikimedia.org/wiki/Accueil
- http://meta.wikimedia.org/wiki/Aide:Liste_de_suivi
- http://meta.wikimedia.org/wiki/IRC_channels
- http://meta.wikimedia.org/wiki/Main_Page
- http://meta.wikimedia.org/wiki/MediaWiki
- http://meta.wikimedia.org/wiki/Wiki_Research_Bibliography
- http://meta.wikimedia.org/wiki/Wikimedia_Research_Network
- http://wikimediafoundation.org/wiki/Fundraising
- http://wikimediafoundation.org/wiki/Home
- http://www.wikimedia.de/
- http://www.wikimedia.fr/
- http://www.wikimedia.it/
- http://www.wikimedia.org/

**Wikipedia allemande**

• http://www.amazon.de/exec/obidos/ASIN/3898530205/303-8124302-7961803

**Wikipedia française**

• http://fr.wikipedia.org/wiki/Aide:Page_Utilisateur
• http://fr.wikipedia.org/wiki/Aide:Sommaire
• http://fr.wikipedia.org/wiki/Bac_a_sable
• http://fr.wikipedia.org/wiki/Portail:Accueil
• http://fr.wikipedia.org/wiki/Special:Recentchanges
• http://fr.wikipedia.org/wiki/Special:Specialpages
• http://fr.wikipedia.org/wiki/Wikilove
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Accueil
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Articles_de_qualit%C3%A9
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Bots
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Cat%C3%A9gorie
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Conflits_entre_%C3%A9diteurs
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Coordination_multilangues
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Guerres_d%27%C3%A9dition_en_cours
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Harmonisation_des_messages_d%27avertissement
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:La_neutralit%C3%A9_de_point_de_vue
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Liens_interlangues
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Liste_des_mod%C3%A8les#G%C3%A9ographie
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Liste_des_mod%C3%A8les#Domaine_du_vivant
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Mod%C3%A8les
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Pages_m%C3%A9ta
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Projet/Les_cahiers_de_Wikip%C3%A9dia
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Projet/Traduction
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:R%C3%A9ponses_aux_objections_habituelles
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Recommandations_et_r%C3%A8gles_%C3%A0_suivre
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Travaux_in%C3%A9dits
• http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Vandalisme

**Wikipedia anglaise**

• http://en.wikipedia.org/wiki/Category:Protected
• http://en.wikipedia.org/wiki/Help:Contents
• http://en.wikipedia.org/wiki/List_of_topics_lists
• http://en.wikipedia.org/wiki/Portal:Browse
• http://en.wikipedia.org/wiki/Special:Recentchanges
• http://en.wikipedia.org/wiki/Special:Specialpages
• http://en.wikipedia.org/wiki/Wikipedia:Counter_Vandalism_Unit
• http://en.wikipedia.org/wiki/Wikipedia:No_original_research
• http://en.wikipedia.org/wiki/Wikipedia:Pushing_to_1.0
L'édition de référence libre et collaborative : le cas de Wikipedia