

Journée professionnelle de l'ADBEN Angers : *Wikipédia : comment appréhender cet outil collaboratif avec des élèves ?*

CDDP d'Angers, 31 mai 2008



Laure Endrizzi

laure.endrizzi@inrp.fr

chargée d'études et de recherche,
Veille scientifique et technologique, INRP
<http://www.inrp.fr/vst/>

support distribué sous licence [Creative commons BY-NC](https://creativecommons.org/licenses/by-nc/4.0/)



Wikipédia 2001-2008 : état des lieux

- Wikipédia : une nouvelle utopie encyclopédique ?
- Caractéristiques du projet et travail éditorial
- Présentation des ateliers



Wikipédia 2001-2008 : état des lieux



**Wikipédia dans le paysage
éditorial contemporain
Une nouvelle utopie
encyclopédique**

3

Un projet parmi d'autres



WIKIMEDIA

a multilingual free
encyclopedia

Wiktionary
[ˈwɪkʃənri] n.,
a wiki-based Open
Content dictionary
Wɪʃən [ˈwɪʃən kənri]



Wikiquote



WIKISOURCE



WIKINEWS



WIKIMEDIA
COMMONS



WIKISPECIES
free species directory



WIKIVERSITY

<http://wikimediafoundation.org/>

4

Un produit illimité et multilingue

WIKIPEDIA

English
The Free Encyclopedia
2 172 000+ articles

Deutsch
Die freie Enzyklopädie
691 000+ Artikel

Français
L'encyclopédie libre
606 000+ articles

Polski
Wolna encyklopedia
459 000+ hasel



Português
A enciclopédia livre
352 000+ artigos

Nederlands
De vrije encyclopedie
398 000+ artikelen

Italiano
L'enciclopedia libera
395 000+ voci

Svenska
Den fria encyklopedin
269 000+ artiklar

<http://www.wikipedia.org/>

Plus de 200 000 articles dans une 12aine de langues

Un espace inédit d'accès à la connaissance mondiale

- **Des modes d'accès hybrides et centrifuges**
 - Intégration au poste de travail :
 - extensions intégrées au navigateur web
 - raccourcis de recherche sur Google et Yahoo
 - raccourcis dans les ENT et campus numériques
 - Accès distant via des outils de recherche variés
 - moteurs généralistes : Google, Yahoo et MSN
 - moteurs alternatifs tels qu'Exalead, Factbites et Clusty
 - métamoteurs d'ouvrages de référence : Answers et Onelook
 - métamoteurs pour wikis tels que WikiSeek ou Qwika
 - Accès distant via des outils mobiles
 - interfaces pour téléphone portable ou assistant personnel

Via Exalead par exemple...

The screenshot shows the Exalead search engine interface. The search term is 'bibliothèque', and it has returned 8,037 results. The main results list includes:

- Bibliothèque**: Une bibliothèque - du grec bibliothéké, lieu de dépôt de livres - est une collection organisée de livres, généralement accessible au public. Les bibliothèques proposent ...
- Ancienne bibliothèque centrale de Montréal**: L'ancienne bibliothèque centrale de Montréal était la bibliothèque principale de Montréal avant la Grande bibliothèque du Québec. L'édifice est situé au 1210 rue ...
- Bibliothèque du Temple d'Auguste**: La bibliothèque du Temple d'Auguste (Bibliotheca Templi Augusti ou Bibliotheca Novi Augusti) fut établie par l'empereur Tibère et consacrée après sa mort. Il avait ...
- Bibliothèque nationale de France**: La Bibliothèque nationale de France (BnF) est la plus importante bibliothèque de France. Elle a le statut d'établissement public. Ses activités sont réparties sur ...

On the right side, there is a sidebar titled 'Préciser la recherche' (Refine search) with a list of related terms and categories such as 'Résultats en Français', 'Académie des inscriptions et belles-lettres', 'Arsenal Austerlitz Bibliothèque américaine', 'Bibliothèque logicielle', 'Bibliothèque municipale', 'Bibliothèque numérique', 'Bibliothèque parisienne', 'Bibliothèque publique', 'Bibliothèque spécialisée', 'Cain', 'Nom de Bibliothèque', 'Python', and 'Sciences de l'information'.

Wikipédia et Google se renforcent mutuellement

Analyse du trafic entrant

The chart, titled 'Wikipedia's Upstream Traffic Sources', shows the percentage of traffic coming from various sources over time. The y-axis represents the percentage from 0% to 50.0%. The x-axis shows dates from February 11, 2007, to February 10, 2008. The data points are:

- WWW.GOOGLE.COM: 49.57%
- SEARCH.YAHOO.COM: 11.04%
- WWW.MYSPACE.COM: 4.03%
- SEARCH.MSN.COM: 2.89%

Other sources shown in the legend include WWW.ASK.COM and WWW.MYSPACE.COM. The chart includes a watermark for 'hitwise Real-time competitive intelligence'.

Google = 50% du trafic entrant sur WP

Google + Yahoo + MSN = 65% du trafic entrant sur WP

WP = 2% du trafic sortant de Google

Prescott, LeeAnn (février 2007). *Google Traffic To Wikipedia up 166% Year over Year*. Hitwise. [en ligne] http://weblogs.hitwise.com/leeann-prescott/2007/02/wikipedia_traffic_sources.html

Wikipédia et Google, référencement et modèle économique

- Parenté Google et Wikipédia :
 - vocation complémentaire dans l'accès et l'organisation de la connaissance du web mondial
 - 2 logiques qui s'articulent sans transaction financière
- Pagerank de WP solide :
 - fondé sur une multiplicité de sites et de réseaux sociaux (blogs et sites web 2.0)
 - et sur un maillage hypertextuel interne stable et serré
- Modèle économique hybride, garanti par la Fondation Wikimedia, qui combine bénévolat des participants et gratuité des contenus

9

Un projet éditorial singulier

- Un objectif commun - une encyclopédie libre et ouverte à tous - et 5 piliers dont la NPOV
- Un état d'inachèvement permanent
 - des questionnements spécifiques en terme de stabilisation et de sécurisation des contenus
 - pertinence encyclopédique négociée et réinventée de façon continue partout (biais culturels)
- Absence de modèle éditorial ? ou encyclopédisme d'usage (O. Ertzscheid) ?
 - WP réunit l'ensemble des savoirs de notre expérience culturelle quotidienne...
 - WP représente une nouvelle utopie encyclopédique,

10

1/3 des internautes français 1er site non commercial dans le top 10

- **Baromètre Médiamétrie** (février 2007)
 - 10e rang des sites français les plus visités en janvier 2007
 - près de **8 millions** de visiteurs uniques : +161% en un an
 - soit 1 internaute français sur 3, dont 60% sont âgés de plus de 25 ans
- **Baromètre ComScore** - octobre 2007
 - 10e rang des sites français les plus visités en juillet / août 2007, près de **7,5 millions** de visiteurs uniques (28% des internautes français)
 - Après Google sites, Microsoft sites, France Telecom, Iliad/Free, eBay, Yahoo! Sites, Groupe Pages Jaunes, Skyrock network, Groupe PPR
 - Et aussi :
 - 10e rang des sites UK, 11,5 millions de visiteurs uniques (36%)
 - 6e rang des sites allemands, 12,8 millions de visiteurs uniques (39%)

11

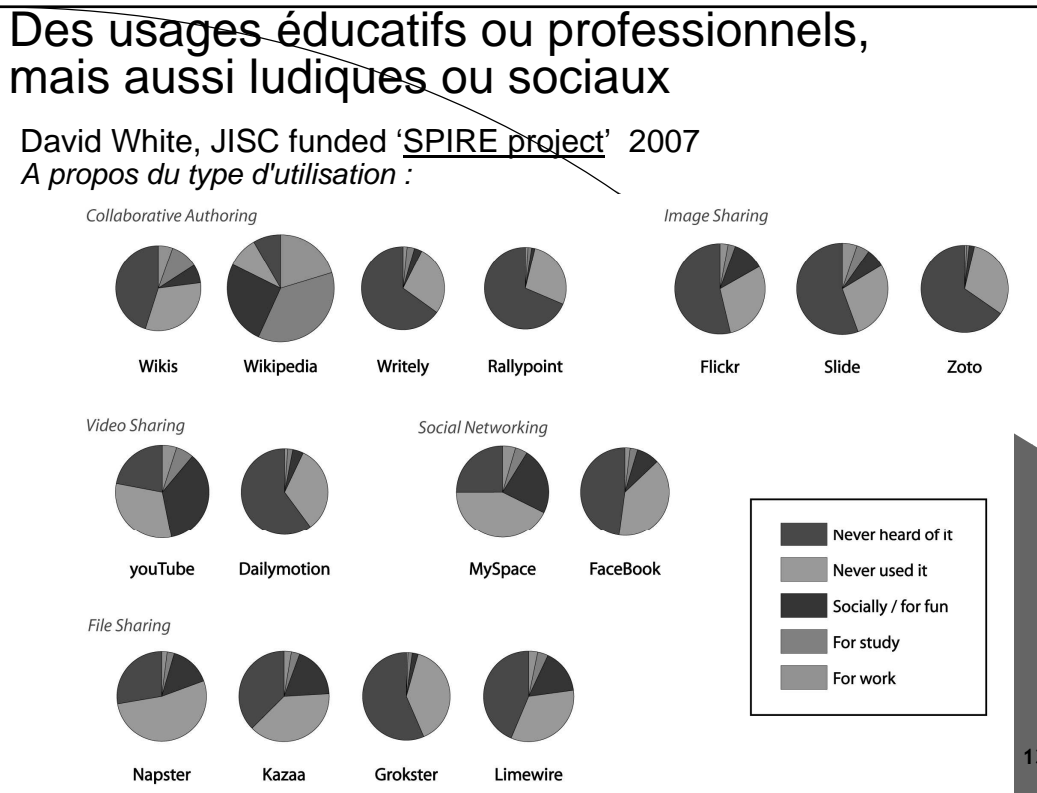
Un public multi-générationnel

A propos du niveau d'utilisation :

- près de 80% des internautes interrogés consultent Wikipédia
- le taux est plus élevé chez les moins de 19 ans et les plus de 65 ans

White David (juin 2007). *Results and analysis of the Web 2.0 services survey undertaken by the SPIRE project*. Oxford : University of Oxford. [en ligne]
http://spire.conted.ox.ac.uk/trac_images/spire/SPIRESurvey.pdf

12



13

Des usagers parmi les privilégiés

36% des internautes américains adultes consultent Wikipédia. Ce taux est plus élevé chez les internautes

- âgés de 18 à 29 ans
- diplômés de l'enseignement supérieur (après licence)
- dont le niveau de vie et le pouvoir d'achat sont importants

WP s'inscrit de plus en plus dans les usages ordinaires d'internet : plus « quotidiens » que la consultation de sites marchands ou de sites de rencontres...

Cet élargissement touche d'abord les populations culturellement et économiquement privilégiées.

A Profile of Wikipedia Users	
Do you ever use the internet to look for information on Wikipedia?	% of adult internet users who say 'yes'
Men	39%
Women	34%
Whites	37%
Blacks	27%
English-speaking Hispanics	36%
Age 18-29	44%
Age 30-49	36%
Age 50-64	31%
Age 65+	26%
High school diploma	22%
Some college	36%
College grad +	50%
< \$30,000 household income	32%
\$30,000-\$50,000	35%
\$50,000-\$75,000	39%
>\$75,000	42%
Dial-up connection at home	26%
Broadband at home	42%

Source: Pew Internet & American Life Project Survey Feb. 15-March 7, 2007. N=1,492 Internet users. Margin of error is ±3%.

Rainie Lee & Tancer Bill (avril 2007). *Wikipedia users*. Washington : Pew Internet & American Life Project. [en ligne] http://www.pewinternet.org/pdfs/PIP_Wikipedia07.pdf

14

Web 2.0 et culture participative

- Une évolution technologique, mais surtout culturelle, qui affecte toutes les strates de la production / diffusion d'information (grand public, professionnelle, scientifique)
 - Journalisme participatif, citoyen
 - Documentation : projets Gutenberg et ODP
 - Cours en ligne (MIT), archives ouvertes, manuels Sésamath, etc.
- “ Web 2.0 ” = read / write web de Tim Berners Lee, conçu comme un outil de collaboration pour les scientifiques du CERN
- lecteur aujourd'hui, indexeur et auteur demain ?
 - Banalisation des sites web 2.0 et émergence des pratiques contributives
 - Avec WP, les moins de 25 ans sont majoritairement des lecteurs ; les contributeurs se recrutent davantage parmi les plus de 45 ans

15

Succès de Wikipédia aujourd'hui

- WP constitue un ensemble organisé de savoirs,
 - couvrant potentiellement tous les sujets (profanes ou savants), au milieu de la jungle informationnelle du web
- WP s'appuie sur une forte corrélation avec les moteurs de recherche
 - Forte popularité d'outils tels que Google, Yahoo, MSN
 - Excellent classement de WP dans les résultats, dû à son « maillage hypertextuel » (interne et externe)
- WP est « pratique » : accessible de n'importe où, n'importe quand
 - Pour les usagers, la « convenance » importe plus que l' « accuracy » : 71% contre 13%

Rainie Lee & Tancer Bill (avril 2007). *Wikipedia users*. Washington : Pew Internet & American Life Project. [en ligne] http://www.pewinternet.org/pdfs/PIP_Wikipedia07.pdf

16

Wikipédia 2001-2008 : état des lieux



**Caractéristiques du projet
et travail éditorial**
**stabilisation et sécurisation de l'espace
éditorial**

17

Principes de base

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




- ***Un objectif commun : rédaction d'une encyclopédie, généraliste et spécialisée, multilingue et gratuite***
- **Impartialité et neutralité de point de vue (NPOV) :**
 - Agrégation de tous les points de vue significatifs (≠ « représentation du point de vue le plus courant »)
- **Propriété intellectuelle et licence GFDL :**
 - Consultation libre, modification libre, distribution libre
 - Commercialisation autorisée sous condition
 - Bénévolat et rédaction collective, articles non signés

18

Principes de base

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



- **Savoir vivre, WikiLove**
 - Respect des opinions et des cultures
 - Respect des personnes (civilité, politesse)
 - Respect des tâches de maintenance et bonnes pratiques
- **Audace, pas de règles fixes : « be bold »**
 - Culture de l'inachèvement
 - Droit à l'erreur

19

Wiki et transparence des mécanismes gestionnaires

- Wiki : concept théorisé par Ward Cunningham et expérimenté pour la 1ère fois en 1995
- Aujourd'hui, des dizaines de moteurs wiki open source, aux fonctionnalités plus ou moins distinctives, mais ce sont tous...
- ... des environnements collaboratifs qui favorisent
 - l'auto-gestion,
 - la coopération décentralisée,
 - la régulation a posteriori des contenus (workflow, versioning),
- ... des environnements collaboratifs où
 - contenus, méta-informations et outils de communication sont intimement corrélés, au service d'une activité réflexive et critique explicite

20

Wiki et transparence des mécanismes gestionnaires

- Visibilité et interdépendance des mécanismes gestionnaires :
 - Traçabilité de la rédaction (versioning)
 - Traçabilité du travail éditorial (gestion, règles)
 - Traçabilité de l'évolution de la communauté
 - Une sorte de « micro-organisme »
- Autrement dit, les pages articles ne représentent qu'une partie des contenus :
 - Entre 60% et 80% des contenus selon les encyclopédies
 - Article : une cellule de base autour de laquelle d'autres cellules viennent se greffer
 - Plus l'encyclopédie est jeune, plus les cellules de type « article » sont proportionnellement dominantes

21

Wiki et transparence des mécanismes gestionnaires

- Les espaces auxiliaires (namespaces)
 - Fichiers multimédia (images, sons, vidéos), Pages utilisateurs, Pages portails, Pages projets,...
 - Pages aide, Pages modèles, Pages catégories, Pages Wikipedia (méta), Pages MediaWiki,...
 - Et pour chacun de ces espaces, les pages de discussion correspondantes
- Des modes d'appropriation **variables et évolutifs selon les univers linguistico-culturels et la maturité éditoriale**
- Exemple de la WP EN entre 2003 et 2005 :
 - Pages de discussion des utilisateurs x 78
 - Pages Wikipedia (méta) dédiées à la gestion du projet x 68
 - Multiplication des wikiprojets et des catégories « ébauches » (stubs)
 - Pages de discussion des articles de plus en plus utilisées pour la coordination, la planification et la dynamisation de l'activité éditoriale

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

22

Acculturation des wikipédiens

- A l'origine, une ritualisation de l'activité :
 - ◊ les rôles se diversifient tout en se spécialisant => émergence de formes d'expertise diversifiées
- une évolution des représentations
 - ◊ d'une collection d'articles à une collection d'individus => micro-conversations
- une croissance qui se joue au niveau local
 - ◊ Parainage informel de la part des contributeurs les plus engagés, donc les plus visibles => accueil et parainage formels (comité d'accueil et service de parainage actif)
- incidence réduite de la biographie extérieure au projet encyclopédique sur la reconnaissance interne

Bryant, Susan L., Forte, Andrea & Bruckman, Amy (2005). Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. In *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*. New York : ACM Press.
 Capocci A., Servedio V., Colaioni F., « Preferential Attachment in the Growth of Social Networks: The Internet Encyclopedia Wikipedia ». *Physical Review E*, vol. 74, n°3, 2006.

23

Acculturation des wikipédiens

- Les ingrédients de l'acculturation, des motivations multiples :
 - ◊ Jeu et plaisir : « edutainment »
 - ◊ Adhésion au projet communautaire, sentiment de participer à un bien commun informationnel
 - + une certaine culture du don = relative discrétion sur l'activité
 - + quête de reconnaissance (voire carriérisme) = publicisation des activités
 - ◊ Importance clé de la biographie personnelle dans les niveaux et stratégies d'engagement

- Bryant, Susan L., Forte, Andrea & Bruckman, Amy (2005). Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. In *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*. New York : ACM Press.
 - Levrel, Julien (octobre 2005). *Développement des contenus libres sur Internet - Wikipedia : approche de l'organisation en ligne et motivations des contributeurs*. France Telecom R&D TECH/SUSI.

24

Formes de l'expertise gestionnaire

- Des statuts qui dessinent une hiérarchie de pouvoirs techniques
 - pour protéger des pages, restaurer une version, bloquer des contributeurs malveillants, supprimer des pages ou des images sous copyright,...
 - répondre à une demande légale d'identification, retirer une version diffamatoire (oversights), ...
 - distribués à des candidats ayant « fait leurs preuves »
- Dans la WPFR - liste consultée en déc. 2007 :
 - 165 administrateurs, 7 bureaucrates, 5 vérificateurs d'utilisateurs, 1 oversight, 3 développeurs
 - soit moins de 1% des contributeurs enregistrés

25

Formes de l'expertise gestionnaire

- Des fonctions plus ou moins formelles en émergence :
 - ♦ Patrouilleurs RC, WikiPompiers, WikiLovers
 - ♦ WikiFourmis, WikiGnômes, WikiTraducteurs
 - ♦ WikiFées, WikiPhotographes, WikiGraphistes, etc.
 - ♦ Lutte contre le vandalisme et les dérives commerciales
- Et des styles d'engagements diversifiés, plus pragmatiques qu'idéologiques, et motivés par le plaisir de participer et d'apprendre...

À découvrir sur la WP FR

- Aide: les différents rôles et Wikipédia: statuts des utilisateurs
- Wikipédia: répartition des Wikipédiens : par localisation géographique, par domaine de compétences, par année de naissance, par animal domestique, par heure de lever,...

26

Formes de l'expertise éditoriale

- **Dynamisation = publicisation des besoins**
 - Wikiprojets, portails thématiques, focus temporaire
 - Catégorisation (ébauches) et application de modèles liés au cycle de vie de l'article
 - Rôle des médias et articles des détracteurs : « effet piranha »
- **Normalisation de la structure, de la navigation et indexation des articles** (modèles et catégories thématiques)
- **Gestion des conflits d'édition et arbitrage** (wikipompiers et CAr)
- **“Validation” : plus ou moins formalisée**
 - Suppression des pages pauvres ou non pertinentes : PàS (WP EN)

Riehl D. (2006) « Les articles de qualité : AdQ (WP DE) ». In *WikiSym '06: Proceedings of the 2006 international symposium on Wikis*. New York : ACM Press, p. 3–8.

27

Formes de l'expertise éditoriale

- **Raffinement progressif des règles et conventions, variables selon les versions**
 - Importance de la vérifiabilité
 - exemples de retrait du label AdQ
 - exemples d'arbitrage sur la “long term notability” pour gérer l'auto-promotion
- **Évaluation et stabilisation, via notamment Wikipédia 1.0**
 - Près de 900 articles sélectionnés au niveau des projets pour leur pertinence encyclopédique et leur maturité
- **Automatiser la visibilité des besoins et la répartition des tâches : une priorité**

Stvilia B., Twidale M.B., Gasser L., Smith L.C., *Information Quality Discussions in Wikipedia*. Champaign, University of Illinois at Urbana-Champaign, 2005

Forte A., Bruckman A., « Scaling Consensus: Increasing Decentralization in Wikipedia Governance ». In *HICSS '08: Proceedings of the Proceedings of the 41st Annual Hawaii International Conference on System Sciences*. Washington DC, IEEE Computer Society, 2008.

Viégas F.B., Wattenberg M., McKeon M.M., « The Hidden Order of Wikipedia ». In *OCSC 2007: Proceedings of the Second International Conference on Online Communities and Social Computing*. Springer, 2007, pp. 445-454

28

En débat : qui écrit Wikipédia ?

- des auteurs occasionnels, anonymes ou non, extérieurs à la communauté, motivés par le projet
 - En filigrane, un modèle éditorial classique :
 - Des outsiders - majoritaires - qui fournissent le contenu
 - Des insiders - minoritaires - qui le corrigent, le reformatent, l'indexent, et gèrent les auteurs
- Ou bien des auteurs fortement impliqués, des contributeurs actifs, voire hyperactifs, qui comptabilisent le plus d'interventions ?
- Dans tous les cas, plus de contributeurs = plus de qualité

Swartz A., « Who writes Wikipedia ? » *Raw Thought*. 2006 [en ligne]

<http://www.aaronsw.com/weblog/whowriteswikipedia>

Kittur A., Suh B., Pendleton B.A., Chi E.H., « He Says, She Says: Conflict and Coordination in Wikipedia ». In *CHI '07 : Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, ACM Press, 2007, p. 453-462.

Wilkinson D.M., Huberman B.A., « Cooperation and Quality in Wikipedia ». In *WikiSym '07: Proceedings of the 2007 International Symposium on Wikis*. New York, ACM Press, 2007, p. 157-164

29

En débat : qui écrit Wikipédia ?

- Inégalités fortes entre les contributeurs (et les articles)
- Analyses qualitatives nécessairement partielles
- Métriques qui s'affinent progressivement
 - importance du facteur temps : durée et régularité de l'engagement
 - Importance de la popularité des articles : pages vues (ou pas)
- Identification de routines de travail chez les administrateurs :
 - enchaînement systématique de tâches répétitives (ex. catégorisation, correction ortho, etc.)
 - activités réactives : dépendantes de l'intervention d'un tiers (ex. surveillance modifications récentes)

Ortega F., Gonzalez-Barahona J.M., Robles G., « On the Inequality of Contributions to Wikipedia ». In *HICSS '08: Proceedings of the Proceedings of the 41st Annual Hawaii International Conference on System Sciences*. Washington DC, IEEE Computer Society, 2008

Priedhorsky R., Chen J., Lam S.K., « Creating, Destroying, and Restoring Value in Wikipedia ». *2007 International Conference on Supporting Group Work*, 4-7 novembre 2007.

Wattenberg M., Viégas F.B., Hollenbach K.J., « Visualizing Activity on Wikipedia with Chromograms ». In *Human-Computer Interaction – INTERACT 2007 (part II)*. Springer, 2007, p. 272-287

30

Anonymat / pseudonymat : suspect ?

- Proportion des éditions anonymes variable selon les versions, mais peu évolutive
 - 19% WP FR contre 48% WP JP (déc. 2007)
- Des contributeurs anonymes « bons samaritains » :
 - Peu enclins à faire carrière et à s'identifier à une communauté mais adhérents au projet, intervenant ponctuellement
- Pas de relation établie entre anonymat et vandalisme
- Indices d'un comportement anonyme suspect :
 - Quand la fréquence des interventions est élevée, en particulier dans les pages de discussion
- Question des identités multiples ou “faux nez”

Anthony Denise, Smith Sean W. & Williamson Tim (2005). *Explaining Quality in Internet Collective Goods : Zealots and Good Samaritans in the Case of Wikipedia*. Hanover : Dartmouth College.
 Kittur A., Suh B., Pendleton B.A., Chi E.H., « He Says, She Says: Conflict and Coordination in Wikipedia ». In *CHI '07 : Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, ACM Press, 2007, p. 453-462

Quelle autorité éditoriale ?

- Quels arbitrages : formels, factuels ou véritablement éditoriaux ?
 - Qu'est-ce qui est « encyclopédique » ? anecdotique ?
 - Emphase sur l'actualité, l'événementiel et les “passions” adolescentes
 - Goût immodéré pour les listes et fonction almanach
 - Risque d'une bureaucratisation des processus ?
- Quelles formes d'autorité dans les arbitrages ?
 - Moyenne des points de vue plus pertinente que le point de vue du meilleur des experts ? (cf. théorie sur l'intelligence collective)
 - NPOV et consensus : des notions ambivalentes ?
 - Groupes de pression : les plus persévérants « gagnent », les plus sincères se découragent ?
 - Quelles collaborations entre amateurs et pros ?
 - Rôle des experts dans la régulation de l'intelligence collective ?

32

Wikipédia discréditée par son (absence de) projet éditorial ?

- Le projet éditorial de Wikipédia est-il crédible ?
 - ♦ Un projet encyclopédique ou un produit éditorial à part entière ?
 - ♦ Une encyclopédie des savoirs ? Une encyclopédie de l'usage ? Une encyclopédie des cultures populaires, mêlant savoirs et croyances ?
 - ♦ Des contenus instables dont la fiabilité est forcément relative
 - ♦ Quelle(s) assurance(s) qualité dans un processus d'écriture collaborative permanent ? Puisqu'il ne s'agit pas d'une question d'autorité au sens traditionnel (qualité des auteurs),
 - S'agit-il d'une question de nombre ? ou une question de temps ? ou une question de stabilité ?

Mahé, Annaïg (octobre 2007). Peut-on se fier à Wikipédia ? *Pour la science*, n°360.

33

Savoirs, informations, croyances...

- Les biais culturels sont réels (étude sur WP EN) :
 - Religion : thème populaire (souvent cité) dans la WP EN
 - ancrage fort des articles relatifs au temps et à l'espace dans l'actualité et un recentrage net sur les EU
- L'objet encyclopédique soumis au risque de l'information ou de la croyance :
 - Absence de légitimité culturelle ? nouvelle légitimité culturelle ?
 - Tous les sujets peuvent-ils être « open source » (ex. des sujets historiques) ?
 - Comment lutter contre les risques de dérives sectaires, politiques, marchandes ?

- Bellomi Francesco & Bonato Roberto (2005). « Network Analysis for Wikipedia ». In *Wikimania 2005*, Frankfurt, Germany, août 2005.
 - Legendre Bertrand (2000). « L'encyclopédie numérique au risque de l'information ». *Les Cahiers du numérique*, vol. 1, n° 5, p. 71-81.
 - Rosenzweig Roy (2006). « Can History be Open Source? Wikipedia and the Future of the Past ». *Journal of American History*, vol. 93, n° 1, p. 117-146.

34

Qualité des articles : une question de stabilité ?

- Difficultés liées à la stabilité des contenus, plus qu'à leur qualité :
 - Gestion des liens externes (sources) : 200 par minute dans la version anglaise (déc. 2006)
 - Spamdexing : inflation de liens commerciaux dans les références externes => introduction de balises « no follow »
 - Sources inappropriées : sites web 2.0 par ex (MySpace, YouTube...) => listes noires
 - Contrôle de notoriété : articles surveillés par des groupes de pression (stés commerciales, partis ou hommes politiques, etc.)
 - Recherche de notoriété : auto-promotion dans biographies personnelles (« annuaire »)

35

Qualité des articles : une question de temps ?

- Approches quantitatives
 - Nombre d'éditeurs et de contributeurs uniques par article ?
 - Progression parallèle du « poids » des articles et de la création de nouveaux articles ?
 - Taux de croissance des contributeurs actifs ?
 - Taux de croissance des espaces auxiliaires ?
 - Taux de croissance des articles de qualité ?
 - Nombre de mots ajoutés (et retenus) à chaque édition par rapport au nombre de mots total et au temps d'existence de l'article ?
 - etc.

36

Qualité des articles : une question de temps ?

- Approches qualitatives
 - Indices d'une maturité progressive de l'activité éditoriale
 - Exemple du processus de labellisation des articles de qualité :
 - Critères : degré de précision, exactitude, pluralité des perspectives
 - Et aussi degré de pertinence, référence à des sources externes, cohérence stylistique
 - Scores proches de zéro pour des critères tels que la validation scientifique et la stabilité
 - Évolutivité des critères (FARC)
- Un différentiel qualité peu significatif sur le plan factuel (comparaison avec encyclopédies traditionnelles)

- Stvilia Besiki, Twidale Michael B., Gasser Les & Smith Linda C. (2005). *Information Quality Discussions in Wikipedia*. Champaign : University of Illinois at Urbana-Champaign, [working paper].

- Giles Jim (2005). « Internet encyclopaedias go head to head ». *Nature*, n° 438, p. 900–901.

- Rosenzweig Roy (2006). « Can History be Open Source? Wikipedia and the Future of the Past ». *Journal of American History*, vol. 93, n° 1, p. 117–146.

37

Qualité des articles : une question de réputation ?

Trust coloring tool : le détecteur de réputation

contributions modifiées, corrigées, effacées = contributeur suspect

occasionally recruited from outside the Folketing.
 Since 27 November 2001, the economist Anders Fjogh
 Rasmussen has been Prime Minister to Denmark.
 As known in other parliamentary systems of government,

Figure 1: An example of coloring words by trust, from the Wikipedia article on *Politics of Denmark*, after the prime minister's middle name has been changed from "Fogh" to "Fjogh" by a low reputation user (revision id: 77692452).

38

Quelle assurance qualité ?

- Question du « recrutement » des contributeurs
 - Accueil et acculturation des nouveaux
 - Équilibre à maintenir entre novices et experts
 - Nécessité d'outils d'aide à la décision pour le partage des tâches
- Pari sur la survivance des contenus
 - Outils pour la construction d'une autorité dynamique, plus ou moins héritée de métrique telle que le « facteur d'impact »
- Objectifs prioritaires centrés sur la sécurisation et la stabilisation de l'espace éditorial

39

Entre rupture et continuité, un modèle éditorial hybride

- Des formes d'expertise gestionnaire interne explicite, consensuelle, efficace
 - Pour corriger, mettre en page, enrichir,...
- Des formes d'expertise éditorial interne, appuyée sur des pouvoirs techniques, plus controversée
 - Pour évaluer les besoins, normaliser, indexer, valider les articles
 - Pour gérer les auteurs (et les conflits entre auteurs)
 - Pour définir les règles (et les faire appliquer)
- Des auteurs multiples, contributeurs occasionnels ou actifs, mais émergence d'une métrique de la réputation

40

Entre rupture et continuité, un modèle éditorial hybride

- Jeu d'influences et d'emprunts réciproques avec l'édition traditionnelle
 - Dispositifs d'ouverture : Encarta, Encyclopédie de l'Agora, Encyclopédie Larousse,...
 - Distribution DVD et version imprimée en chantier pour la WP DE
- Une réaffirmation des logiques « a posteriori »
 - « validation » a posteriori des contenus (AdQ et Pàs)
 - Autorité / réputation construite dynamiquement, via un examen automatique des contributions (et de leur pérennité)
- Culture participative : le wiki va se dissoudre dans le web de demain ?

41

Merci de votre attention...



- *Endrizzi L. 2007. La communauté comme auteur et éditeur : Le cas de Wikipédia. In Journée nationale du réseau des URFIST, Evaluation et validation de l'information sur internet, Paris, 31 janvier 2007.*
- *Endrizzi, L. 2008. Wikipédia : un nouveau modèle éditorial ?. In Schöpfel J. (dir), Les publications scientifiques : analyses et perspectives, coll. Traité « STI - Sciences et Techniques de l'Information ». Paris : Hermès. (à paraître).*

Espace wiki de mutualisation :

<http://fr.wikibooks.org/wiki/Wiki-hop>



Laure Endrizzi – laure.endrizzi@inrp.fr

42