

Wikipédia : outil d'élaboration/gestion

des connaissances

Mise à jour du 12 mars 2013

Rémi Bachelet

Ce cours est, avec son accord, largement issu [du dossier](#) de Laure Endrizzi, INRP (signalé en bas des diapos concernées)

Version à jour, éditable et animée disponible ici :
[Modèle éditorial de Wikipédia et efficacité de la coopération décentralisée](#)



Cours distribué sous licence **Creative Commons**,
selon les conditions suivantes :



Source des images indiquées au-dessous ou en cliquant sur l'image

Image : [Source](#)

7/ Modèle éditorial et efficacité de la coopération décentralisée

- Quels sont les fondements philosophiques de Wikipédia ?
- Wikipédia est-elle crédible ?
- Quelles sont les formes du vandalisme et les moyens de le combattre ?
- La qualité vient-elle avec le temps ?
- Comment se portent les autres projets « encyclopédiques » sur la Toile (Knol, Citizendium) ?

Fondements philosophiques de Wikipédia

5 Courants pertinents :

- Libéralisme
- Anarchisme
- Relativisme
- Rationalisme
- Positivisme
- .

Fondements philosophiques de Wikipédia

■ Libéralisme

- Tout le monde doit profiter de l'expression, de la circulation et de la critique des opinions de tous.
- Une résistance très forte à la censure
 - Cas des images de Commons

■ Anarchisme

- Capacité des individus à se gouverner collectivement de manière autonome
- Les règles de chaque projet sont décidées par la communauté de ce projet
- Chaque éditeur est responsable de ses actions

Fondements philosophiques de Wikipédia

■ Relativisme

- Tous les points de vue et toutes les opinions ont un droit égal à l'expression.
- Chacun a le droit de contribuer à Wikipédia,
 - Pourvu que ces idées soient reconnues par des références d'experts du sujet et prennent une place proportionnelle à leur importance.

■ Rationalisme

- Le lecteur de Wikipédia, a les capacités de décider parmi les différentes théories ou points de vue ce qu'il doit penser

■ Positivismisme

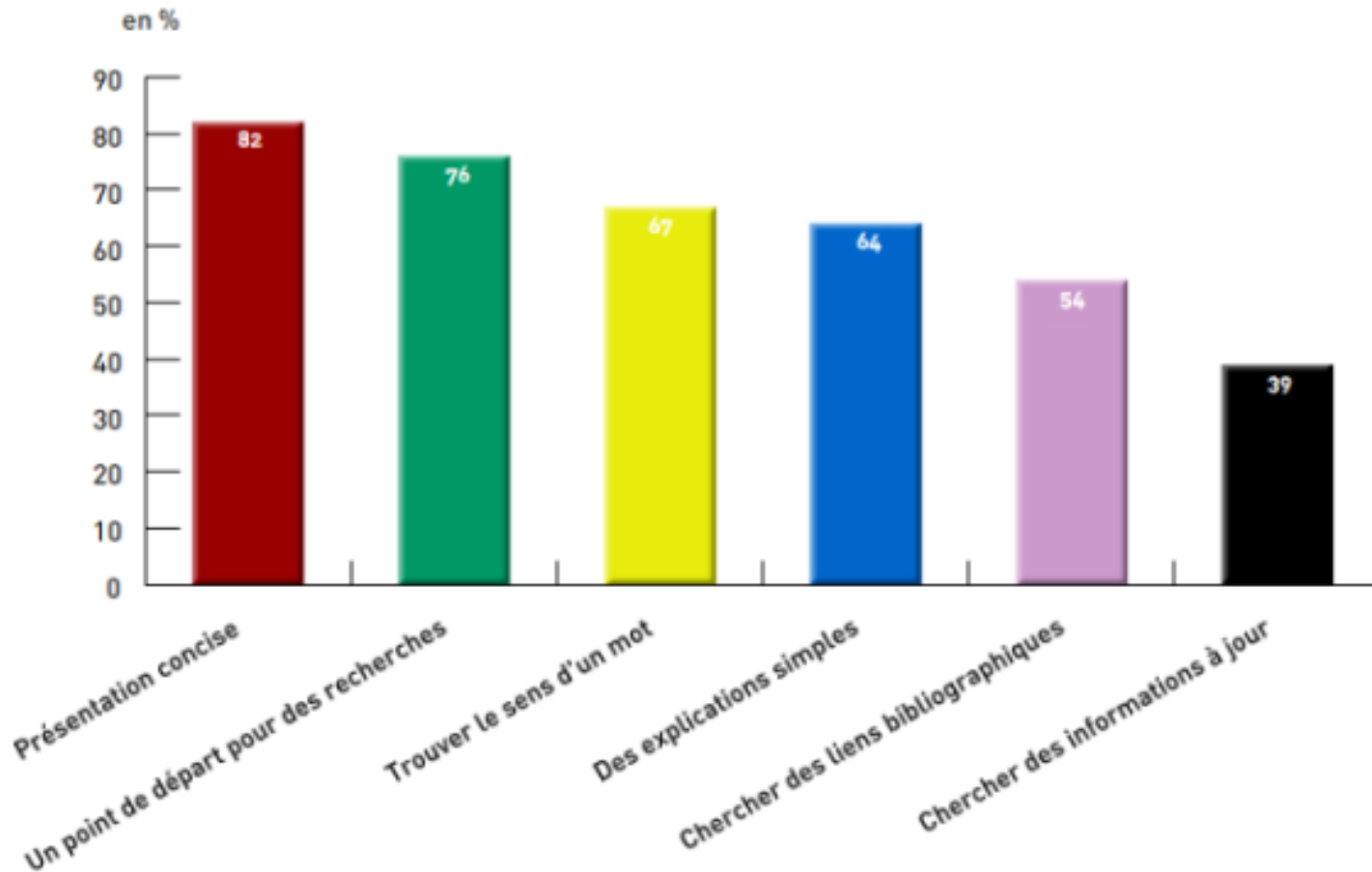
- Exige la vérifiabilité : des sources et des références :
 - du point de vue des contenus, Wikipédia n'est pas une expérience démocratique. On ne « vote » pas le contenu d'un article, mais on cite des sources secondaires reconnues.

Wikipédia est-elle crédible ?

- Un projet encyclopédique ou un produit éditorial ?
- Une encyclopédie des savoirs ou une encyclopédie des cultures populaires ?
- Des contenus dont la fiabilité est forcément relative
 - Quelle(s) assurance(s) qualité ?
 - dans un processus d'écriture collaborative
 - dans le changement permanent ?
- La qualité d'un article
 - Puisqu'il ne s'agit pas d'une question d'autorité au sens traditionnel (qualité des auteurs), s'agit-il
 - D'une question de temps ?
 - D'une question de nombre d'éditeurs ?
 - D'une question de stabilité ?
 - .

Les principales raisons d'une consultation de Wikipedia

Citées par 2 300 étudiants interrogés par l'université de Washington.



source "Comment les étudiants utilisent-ils Wikipedia ?" LeMonde.fr 17.03.2010

.. Au-delà de la vérité

En réalité, que désire une personne qui tape sur Google le mot-clé qui l'intéresse suivi de « wiki » ?

- La vérité ?
- Des données à jour ?
- Des informations résultant d'un consensus entre plusieurs auteurs ?
- Avoir l'essentiel gratuitement et rapidement pour approfondir ultérieurement ?
- Des choses qu'il ne trouverait pas dans une encyclopédie classique ?
- Lire la page de discussion accolée à l'article ? Parcourir les notions évoquées dans celui-ci (hyperliens) ?
- De l'information "active" = pouvoir corriger ou ajouter « son grain de sel » ?

« L'essor de Wikipédia va de pair avec une évolution générale de notre relation à la connaissance et à la fiabilité ».

Wikipédia : principaux problèmes

- Difficultés liées à la stabilité des contenus, plus qu'à leur qualité :
 - Gestion des liens externes (sources) : 200 par minute dans la version anglaise
 - ✓ Spamdexing : inflation de liens commerciaux dans les références externes
 - balises « <nofollow> »
 - Abuse filter, Robots RC comme [Utilisateur:Salebot](#)
 - ✓ Sources inappropriées : sites web 2.0 par ex (MySpace, shortURL...)
 - listes noires
 - Contrôle de réputation : articles surveillés par des groupes de pression (stés commerciales, partis ou hommes politiques, etc...)
 - Recherche de notoriété (biographies personnelles ou règlements de comptes sur les pages des personnes/idées que l'on n'aime pas, saison des élections).
 - Vandalisme (phénomène rentrée des classes « Kevin C 1 PD »)
 - Plagiat (violation de copyright = « copyvio »)
 - [Wikipédia:Réponses aux objections habituelles](#), [Wikipedia Watch](#), [Wikipedia Review](#)
 - ..

Les critiques (détail)

- les contributeurs quand ils ne se **dissimulent** pas derrière un ou plusieurs pseudonymes, sont des anonymes ; il ne sont juridiquement pas responsables de leurs contributions (**faux**) ;
- 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 ; ils sont au mieux des amateurs, au pire des perturbateurs, des vandales ou des fanatiques ;
- hormis ceux traitant d'informatique, les articles sont au mieux de **qualité médiocre**, ou de simples ébauches (**faux**)
- certains thèmes ou **sujets controversés** sont décrédibilisés par des prises de positions partisans (biographies, événements historiques, mouvements politiques, religion, etc.) ;
- les **sujets d'actualité** sont surreprésentés 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 ;
- les articles ne sont pas systématiquement relus ou validés ;
- le manque de fiabilité et de stabilité des articles les rend inexploitable ;
- le qualificatif « encyclopédie » est scientifiquement inapproprié ;
- 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 n'est pas applicable ;
- la **vérité encyclopédique est « votée »** ; quelle représentativité pour les votants ? (**faux**) ;
- le fondateur de wp est un personnage immoral ;
- la représentativité des univers culturels des locuteurs d'une même langue est biaisée ;
- etc.

Savoir, information, croyances ?

wp comme « baromètre des sujets de société » ?

- Etude sur la wp {en}:
 - Thème de la religion fortement représenté - ancrage fort dans l'actualité - centrage net les EU
 - En fait, approche peu pertinente
 - ✓ Ce qui est remarquable c'est la diversité des sujets traités cf. le phénomène longue traîne (« long tail ») : 3M d'articles
- Des biais culturels et des risques de dérives sectaires, politiques, marchandes :
 - Comment les identifier, les mesurer, lutter contre ?
- L'objet encyclopédique soumis au risque de l'information ou de la croyance ?
 - wp comme encyclopédie des cultures populaires : nouvelle légitimité culturelle ?
- L'objet encyclopédique permet-il de traiter les sujets d'actualité ?
 - Ou faut-il du temps et du recul ?

.

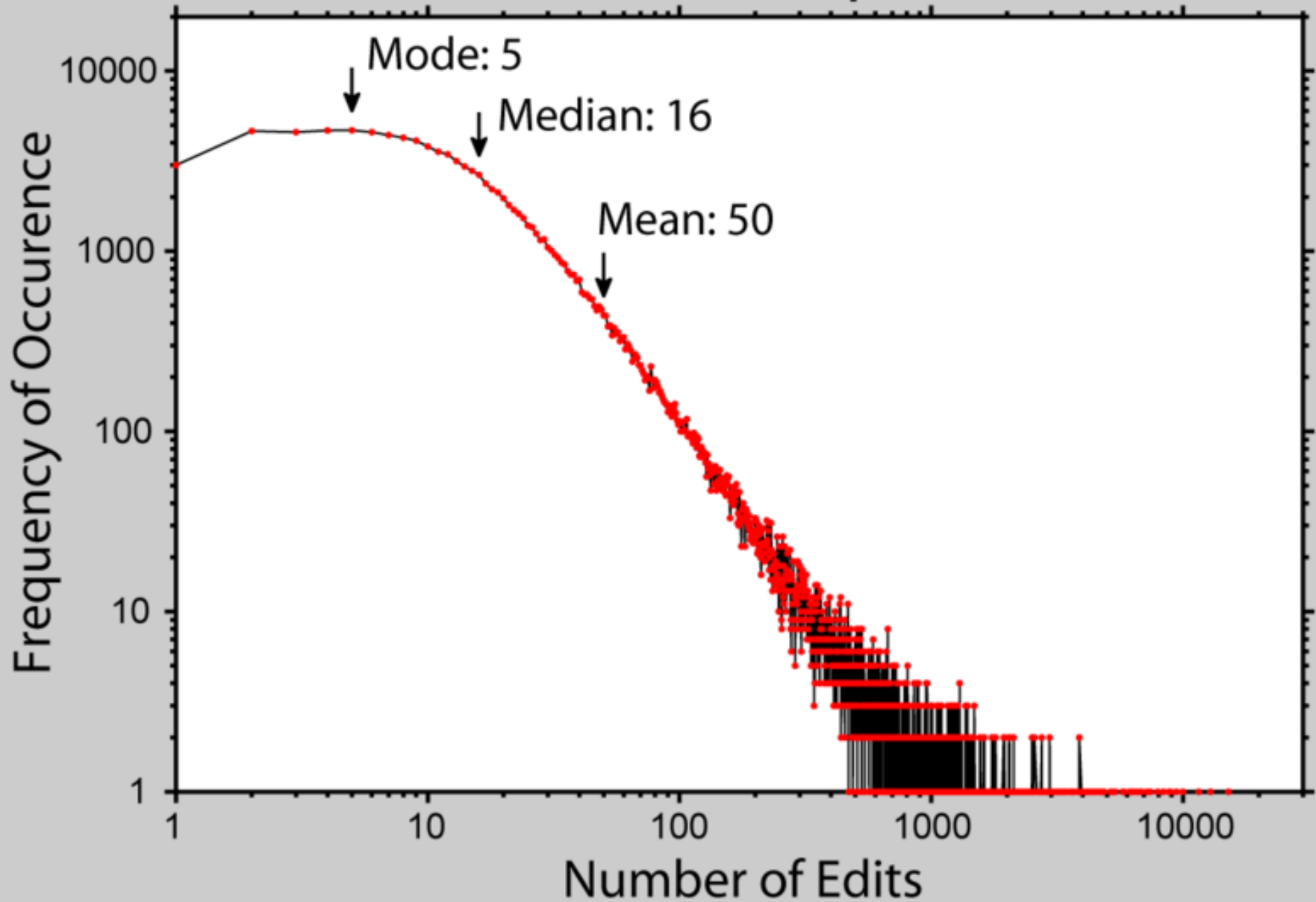
Coopération décentralisée : efficace ?

- Emprunts de wp au Style « bazar » (vs style « cathédrale »)
 - Un artefact simple et approprié
 - Des volontaires nombreux, compétents et motivés
 - Un processus de développement léger et transparent
 - Une gouvernance légère mais efficace, basée sur un continuum d'échanges
 - ✓ Notion de consensus, voire de « consensus clair ».
 - Conversation et réflexivité au cœur de la production / acquisition de connaissances
 - Développement incrémental adapté
- cf. filiations avec les travaux d'E. Raymond sur les communautés de logiciels libres et d'H. Rheingold sur les communautés virtuelles

La qualité : une question de temps ?

- Approches qualitatives
 - Analyse des pages de discussion, des pages des contributeurs et du processus de labellisation des articles de qualité : les critères récurrents
 - ✓ Degré de précision, exactitude, pluralité des perspectives
 - ✓ Et aussi degré de pertinence, référence à des sources externes,
 - ✓ cohérence stylistique
- Maturité progressive de l'activité éditoriale : wp perfectible par itérations successives
- Un différentiel qualité peu significatif sur le plan factuel (comparaison avec encyclopédies traditionnelles)
- Lire « [Wikipédia : les outils de la confiance](#) »

Number of Edits per Article



Typologie du vandalisme

Formes du vandalisme :

- (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 partial, contrevenant à la NPOV.

D'après Fernanda Viégas, Martin Wattenberg et
Kushal Dave 2004
IBM Watson Research Center

Pour une typologie plus récente en français, voir [Utilisateur:Ofol/Typologie du vandalisme](#)
Salebot = robot anti-vandalisme, et [Wikipédia:Prise de décision/AbuseFilter](#)

Quelle autorité éditoriale ?

- Quels arbitrages : formels, factuels ou véritablement éditoriaux ?
 - ◆ Qu'est-ce qui est « encyclopédique » ? anecdotique ? Promotionnel ?
 - ◆ Emphase sur l'actualité, l'événementiel et les modes
 - ◆ 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 de 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 experts ?
 - ◆ ..

Management et expertise politique

- Poids de la Wikimedia Foundation ?
- Poids des associations wikimédiennes et de ses militants ?
- Quels rôles pour les experts dans la régulation de l'intelligence collective ?
 - cf. Citizendium, knol
 - ..

Schéma rédactionnel concret de wp

- Un schéma rédactionnel ayant trois caractéristiques :
 1. Incrémental :
 - ◆ Construction de l'article par touches successives
 2. Décentralisé :
 - ◆ Des contributeurs multiples participant à la rédaction de plusieurs articles d'une collection thématique
 3. Autorité :
 - ◆ Formation de petits groupes collaborant à plusieurs articles, émergence de micro-conversations
 - ◆ ...

Styles participatifs

- Proportion des éditions anonymes variable selon les versions, mais peu évolutive
 - 20% wp FR contre 52% wp JP
- Pas de relation établie entre anonymat et vandalisme,
 -mais suivi prioritaire des contributeurs anonymes et nouveaux (patrouille RC)
- Des contributeurs anonymes « bons samaritains » :
 - Peu enclins à faire carrière et à s'identifier à une communauté mais adhérents au projet
 - Interventions ponctuelles sur des sujets relevant de leur domaine d'expertise

Figures de l'auteur

- Plus généralement des auteurs occasionnels, anonymes ou non, extérieurs à la communauté, motivés par le projet
 - Coût / bénéfice de l'instauration d'une authentification obligatoire ?
- En filigrane, un modèle éditorial classique :
 - ◆ Des outsiders – majoritaires – qui fournissent le contenu
 - ◆ Des insiders – minoritaires – qui le corrigent, le reformatent, l'indexent, le valident, et gèrent les auteurs
- Le noyau actif des wikipédiens :
 - une hiérarchie de fonctions éditoriales en voie de professionnalisation ?
 - ...

Un modèle éditorial hybride

- Prémices d'un modèle éditorial hybride, combinant
 - production libre
 - production collaborative
 - fonctions éditoriales traditionnelles ?
 - ◆ Une expertise **gestionnaire** interne explicite et relativement efficace : « petites mains »
 - ◆ Une expertise **éditoriale** interne moins formalisée, moins explicite, plus controversée
- Des logiques éditoriales complexes, des résultats fragiles qui portent essentiellement sur la wp EN
 - .

Wikipédia *what's next* ?

- Systèmes à réputation : ex. d'Everything2, Knol
 - 1 article = 1 auteur
 - Commentaire/Votes (sanction) des participants => popularité de l'article et de son auteur
 - Filtrage et hiérarchisation des contenus + des contributeurs (droit à publier)
 - Modèle éditorial plus informel, interactions sociales limitées
- Comparaison des niveaux de standardisation et de structuration des articles :
 - Entrées stylistiquement plus homogènes et mieux organisées dans WP
 - Plus le contrôle post-éditorial est ouvertement collaboratif, plus la ligne éditoriale est explicite...
 - ... plus les articles présentent un degré de formalisation linguistique et structurelle fort
 - ..

Autres projets encyclopédiques « rivaux » sur la Toile

- Le modèle éditorial wikipédien n'est
 - ni le modèle « fermé » avec comité de rédaction et ligne éditoriale,
 - ✓ Citizendium, une scission de Wikipedia s'est créée en septembre 2006, pour rétablir un contrôle par un comité d'experts reconnus,
 - ✓ cumule en septembre 2009 10.000 articles (contre 13 millions pour Wikipédia).
 - ✓ 25 éditeurs faisant plus de 100 contributions/mois, 90 éditeurs en mars 2010
 - ni le modèle « ouvert », avec mise en concurrence des thèses solitaires de chacun qui ont eu le plus de succès.
 - ✓ Knol, de Google construit sur le modèle de la concurrence entre rédacteurs individuels
 - ✓ 1.200.000 articles au bout d'un an d'existence, malgré une forte publicité.
 - ✓ Fermeture complète de Knol le 1er octobre 2012. Il sera remplacé par Annotum, « une plateforme de rédaction et de publication académique open source,

Wikipédia est aussi une encyclopédie d'exception en raison de la cohabitation des savoirs experts et communs/amateurs et de la valeur inégale, et non étanche, de la validité de ces contributions laquelle tend à surdévelopper les capacités de la pensée critique chez les utilisateurs.

Aucun autre livre n'a entretenu de telles dispositions sceptiques chez les lecteurs à l'égard de l'autorité et des savoirs, de leurs statuts finis, construits, provisoires.

Wikipédia a rendu l'ère du soupçon définitive.

Apprendre, c'est prendre ses précautions. Sapere Aude.

Questions ?



ILLUSTRATION: OLIVIER BALEZ

Addendums et données

Vandalisme

- 5% des modifications = vandalisme
- sur 4 ans : 2 millions sur 57
- 42% des dommages réparés avant même d'être vus
- quasiment tous supprimés après 15 révisions
- probabilité de lire un article endommagé : 0.0037
- 4 chances /1000 de lire un vandalisme
- reconnu aisément dans 60% des cas
- **2 chances sur 1000 de lire un vandalisme sans vous en rendre compte**

Sur en.wp, source : R Priedhorsky, J Chen, S T K Lam, K Panciera, L Terveen, J Riedl, Creating, Destroying, and Restoring Value in Wikipedia, GROUP'07, 2007

Quelques stats sur les articles...

	(juin)	Nombre d'articles*			Nouveaux / jour (juin 06)
	2004	2005	2006		
1 WP anglaise	270 000	572 000	1 200 000	+ 2 100	
2 WP allemande	100 000	244 000	422 000	+ 530	
3 WP française	36 000	114 000	301 000	+ 720	
4 WP polonaise	28 000	66 000	234 000	+ 340	
5 WP japonaise	55 000	125 000	230 000	+ 280	
Total WP	685 000	1 800 000	4 200 000	+ 8 400	

* Article = 1 lien interne + 200 caractères mini

	Révisions / article	Articles > 2 Ko	Articles de qualité	Total édit. anonymes
1 WP anglaise	34	33%		30%
2 WP allemande	29	40%		30%
3 WP française	19	34%		19%
4 WP japonaise	23	16%		52%
5 WP polonaise	11	14%		18%
Total WP	21	26%		

- En juin 2006, la Wikipédia française, c'était...
 - Presque 13 000 contributeurs enregistrés
(`{en}`: 111 K, total : 214 K)
 - 700 nouveaux contributeurs par mois
(`{en}`: 6 300 ; total : 12 000)
 - 3 500 contributeurs actifs et 630 très actifs
(`{en}`: 38 300 et 4 250 ; total : 67 000 et 9 250)
 - 111 administrateurs, 7 bureaucrates, 5 vérificateurs d'adresses IP – soit 0,85% d'administrateurs
(`{en}`: 1 007 admi, 24 bureaucrates, 15 IP – soit 0,91% d'administrateurs)

Quelques stats sur les contributeurs...

	Nombre de contributeurs			Nouveaux contributeurs / mois			
	(juin)	2004	2005	2006	2004	2005	2006
WP anglaise		11 000	37 000	111 000	1 000	3 300	6 300
WP allemande		6 000	15 000	29 000	700	1 000	1 000
WP française		1 200	4 500	13 000	180	400	700
WP japonaise		1 300	3 400	8 000	150	170	540
WP polonaise		600	2 000	4 700	60	170	250
Total WP		24 000	79 000	214 000	3 000	7 000	12 000

	Contributeurs actifs*		Rôles			
	(juin 2006)	> 5	> 100	Admin	Bureau.	Vérif. IP
WP anglaise		38 300	4 250	1 007	24	15
WP allemande		6 900	980	231	2	3
WP française		3 500	630	111	7	5
WP japonaise		2 500	330			
WP polonaise		1 400	270			
Total WP		67 000	9 250			

Savoir ou information ?

wp comme « baromètre des sujets de société » ?

Avril 2006

	Wpédiens	Article
1	55	Contrat 1ère embauche
2	30	Ségolène Royal
3	28	Union pour un mouvement pop.
4	27	Charles de Gaulle
4	27	Nicolas Sarkozy
5	26	Blog
5	26	Louis XIV

Mai 2006

	Wpédiens	Article
1	41	Wii
2	37	Da Vinci Code
3	33	Lordi
4	31	Projet Maintenance/analyse IP
4	31	Algérie
4	31	Projet Traduction/demandes
4	31	Catastrophe de Tchernobyl

Juillet 2006

	Wpédiens	Article
1	96	Zinedine Zidane
2	79	Coupe du monde de football 2006
3	70	Conflit israëlo-libanais 2006
4	50	Coupe du monde de football
5	46	Marco Materazzi

Juillet 2006

	Wpédiens	Article
1	96	Zinedine Zidane
2	79	Coupe du monde de football 2006
3	70	Conflit israëlo-libanais 2006
4	50	Coupe du monde de football
5	46	Marco Materazzi

Questions ?

Thématiques et questions

- Un wiki comment ça marche ?
- Genèse du projet Wikipédia et Principes fondateurs
- Un succès énorme ? En version diapos animées
- Modes de gouvernance et objectifs de développement. En version diapos animées
- Structuration interne des pages de Wikipédia.
- Rôles et profils des Wikipédiens.
- Modèle éditorial, critiques et efficacité de la coopération décentralisée. En version diapos animées.



Rémi BACHELET

*Enseignant-chercheur,
Ecole Centrale de Lille*

Mon [CV](#) est disponible ici.

Mes principaux cours à Centrale

Gestion de projet, sociologie des organisations, recueil, analyse et traitement de données, prévention du plagiat, module de marchés financiers, cours de qualité et méthodes de résolution de problèmes, établir des cartes conceptuelles, utiliser Wikipédia et CentraleWiki, formation au coaching pédagogique et à l'encadrement