

Wikipédia : de la co-rédaction au co- développement de la communauté

Conférence Doc Soc
Semaine du Document Numérique
Fribourg, Suisse, 20 et 21 septembre 2006
Laure Endrizzi
laure.endrizzi@inrp.fr

Chargée d'études
Veille Scientifique et Technologique,
Institut National de Recherche Pédagogique
<http://www.inrp.fr/vst/>

Avant propos

- Depuis 2003, VST : <http://www.inrp.fr/vst/>
un ensemble de prestations électroniques au service de l'actualité des recherches en éducation
 - Signalement des nouvelles publications
 - Synthèses thématiques dans des lettres mensuelles et dossiers ponctuels
 - Mises en perspective et commentaires de l'actualité dans un blog Ecrans de veille en Education
 - Capitalisation de nos lectures dans une plate-forme collaborative de gestion bibliographique

A l'origine...

- Un projet de recherche sur les encyclopédies en ligne : ENCYCLEN (ACI éducation et formation)
- A la VST : deux « revues de littérature » et une base bibliographique, le tout consultable en ligne :
 - L'édition de référence libre et collaborative : le cas de Wikipédia (mars 06)
<http://www.inrp.fr/vst/Dossiers/Wikipedia/sommaire.htm>
 - Encyclopédisme et savoir : du papier au numérique (avril 06)
http://www.inrp.fr/vst/Dossiers/Savoir_encyclopedique/sommaire.htm
 - Wikibibliographie Encyclen : http://wikindx.inrp.fr/biblio_encyclen/

Aujourd'hui

- Wikipédia :
 - Un produit éditorial mais aussi un projet autour de la poursuite d'un objectif commun : l' « horizon encyclopédique »
 - Un phénomène sans précédent dont la (les) dynamique(s) reste(nt) peu connue(s)
- Quelques éléments de réflexion pour comprendre les modalités d'échanges et le modèle de construction de l'espace social...
 - ... entre production de contenus et développement de la communauté
 - basés sur quelques travaux de recherche récents

Plan de l'intervention

- Quelques statistiques
- La transparence des mécanismes gestionnaires
- La qualité : quels indicateurs ?
- Savoir ou information ?
- La coopération décentralisée : efficace ?
- Les wikipédiens : expertise gestionnaire ou scientifique ?

<http://stats.wikimedia.org/>

En juin 2006, la Wikipédia française, c'est...

- **+ de 301 000 articles dont 91% catégorisés**
(WP en : 1 200 K, total : 4 200 K)
- **Une moyenne de 720 nouveaux articles par jour**
(WP en : 2 100 ; total : 8 400)
- **Une moyenne de 19 éditions par article**
(WP en : 34 ; total : 21)
- **Un « poids » supérieur à 2 Ko pour 34% des articles**
(WP en : 33% ; total : 26%)
- **19% d'éditions anonymes** au total contre 30% dans la WP en et 52% dans la WP jp
- Pages : **73% articles + 7,5% multimedia + 6,5% utilisateurs + 6% catégories + 3,5% modèles + 2,6% projets, etc.**

<http://stats.wikimedia.org/>

- En juin 2006, la Wikipédia française, c'est aussi...
 - **Presque 13 000 contributeurs enregistrés**
(WP en : 111 K, total : 214 K)
 - **700 nouveaux contributeurs par mois**
(WP en : 6 300 ; total : 12 000)
 - **3 500 contributeurs actifs et 630 très actifs**
(WP 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**
(WP en : 1 007 admi, 24 bureaucrates, 15 IP – soit 0,91% d'administrateurs)

Quelques stats sur les articles...

	Nombre d'articles*			Nouveaux / jour (juin 06)
	(juin) 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%		

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	Bureaucr.	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			

Wiki et transparence des mécanismes gestionnaires

- Environnement collaboratif qui favorise la **coopération décentralisée**, l'**auto-gestion** et la **régulation a posteriori** des contenus
- **Visibilité** des mécanismes gestionnaires : traces de la co-construction de l'article et du co-développement de la communauté
- Des outils pour :
 - Suivre l'actualité de la communauté, ses décisions, ses besoins
 - Évaluer son activité interne et son audience externe
 - S'informer sur les règles et conventions de participation
 - Connaître les profils et rôles des contributeurs
 - Assurer le suivi des articles
- **Interdépendance** : articles, méta-informations et outils de communication asynchrones ; **pages articles** : entre 60 et 80% de l'ensemble selon les versions ([Voss](#))

La qualité : une question de temps ?

- Quels indices de la construction temporelle : approches quantitatives
 - Nombre d'éditions et de contributeurs uniques par article ? ([Lih](#))
 - Progression parallèle du « poids » des articles et de la création de nouveaux articles ? ([Levrel](#))
 - Taux de croissance des contributeurs actifs ? ([Levrel](#))
 - Nombre de mots ajoutés à chaque édition par rapport au nombre de mots total et au temps d'existence de l'article ([Anthony](#) et al.)
 - etc.

La qualité : une question de temps ?

- Quels indices de la construction temporelle : approches qualitatives ([Stvilia](#) et al., 2005)
- Analyse des pages de discussion et des pages des contributeurs : les critères récurrents
 - Degré de précision et exactitude, pluralité des perspectives (NPOV)
 - 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
- WP : moyenne des points de vue d'une masse (?) plus pertinente que le point de vue du meilleur expert ?

La qualité : une question de temps ?

- Quels indices de la construction temporelle : approches qualitatives ([Stvilia](#) et al., 2005)
- Analyse du processus de labellisation des articles de qualité
 - Identité des critères pour les FA et les FARC
 - Évolutivité des critères : introduction de la référence à des sources externes en sept. 04 et application rétrospective sur les FA
 - La définition des standards de qualité répond à une maturation progressive de l'activité éditoriale
- WP : perfectible par itérations successives ?

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

Juin 2006

	Wpédiens	Article
1	142	Coupe du monde de football 2006
2	56	Coupe du monde de football
3	45	Ségolène Royal
4	43	Zinedine Zidane
4	43	Raymond Devos
5	37	Equipe de France de football

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

Savoir, information, désinformation ?

WP comme « baromètre des sujets de société » ?

- Etude de [Bellomi](#) et Bonato sur la WP EN :
 - Indice de popularité (pages le plus souvent citées) : articles dominés par le thème de la religion
 - Indice d'autorité (pages de référence) : ancrage fort des articles relatifs au temps et à l'espace dans l'actualité et un recentrage net sur les EU
- Des biais culturels donc et des risques de dérives sectaires, politiques, marchandes : comment les identifier, les mesurer, lutter contre ?
- L'objet encyclopédique soumis au risque de l'information (et donc de la désinformation...) ? ([Legendre](#))

Coopération décentralisée : efficace ?

- Emprunts de WP au Style « bazar » (vs style « cathédrale ») ([Wagner](#))
 - Un artéfact simple et approprié
 - Des volontaires compétents et motivés
 - Un processus de développement léger et transparent
 - Une gouvernance légère mais efficace, basée sur un continuum d'échanges
 - Conversation et réflexivité au coeur de la production / acquisition de connaissances
- cf. filiations avec les travaux d'E. Raymond sur les communautés de logiciels libres et d'H. Rheingold sur les communautés virtuelles

Coopération décentralisée : efficace ?

- Efficacité des systèmes à réputation : ex. d'Everything2 ([Emigh](#) et al.)
 - 1 article = 1 auteur
 - 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

Les Wikipédiens :

expertise gestionnaire ou scientifique ?

- Des statuts qui dessinent une hiérarchie de pouvoirs techniques
- Des styles d'engagements diversifiés, plus pragmatiques qu'idéologiques, basés sur une spécialisation progressive des tâches
- Importance clé de la biographie personnelle dans les niveaux et stratégies d'engagement
- Adhésion au projet communautaire + quête de reconnaissance (voire carriérisme) = publicisation des activités
- 2 études convergentes : celle de [Levrel](#) et celle de [Bryant](#)

Les Wikipédiens :

expertise gestionnaire ou scientifique ?

- Anonymat et expertise scientifique : plus qu'un épiphénomène ? ([Anthony](#) et al.)
 - Des contributeurs anonymes dont la fréquence des interventions est élevée (et dont les contributions sont mineures) : le maillon faible
 - Des contributeurs anonymes dont les interventions sont peu nombreuses mais ciblées, peu fréquentes mais longues : les « bons samaritains »
- Coût / bénéfice de l'instauration d'une authentification obligatoire ?

En guise de conclusion...

- Convergence globale sur les dynamiques internes (et externes), tendues vers le projet encyclopédique
- Premières recherches, premiers jalons...
- ... Mais encore plus de questions que de réponses...
- WP : un objet inédit aujourd'hui pour étudier les communautés virtuelles et la production collaborative et décentralisée de contenus