

Essais de profilage linguistique des discussions Wikipédia

Lydia-Mai Ho-Dac, Veronika Laippala, Céline Poudat and Ludovic
Tanguy

Université de Toulouse – CLLE-ERSS – Université de Turku – TIAS – University of Nice
Sophia Antipolis

Décembre 2016

Plan

- 1 Wikipédia comme objet d'étude
 - Un peu d'histoire
 - Un objet sociologique
 - Une ressource pour le TAL
 - Un corpus d'étude pour la linguistique
- 2 Profilage des discussions WP
- 3 Classification of "Conflicting" vs. "Neutral" Discussions
- 4 A Bottom-Up Approach to Talk Page Profiling
- 5 Conclusion

Un contexte propice

Projet d'encyclopédies ouvertes et libres [Sahut, 2015]

Encyclopédies collective, gratuites, de référence, non soumises au copyright, accessibles sur internet, dont les auteurs seraient des rédacteurs volontaires, avec validation a posteriori des articles

- 1993 : **Interpedia** (R. Gates) participation de tous mais pas d'outil d'écriture collaborative
- 1997 : **Distributed Encyclopedia** (U. Fuchs, futur Wikipédien.de) contre le "poinds croissant de la sphère marchande sur le web"
- 1999 : **GNUpedia**, puis *GNE's Not an Encyclopedia* (R. Stallman, logiciel libre) "garder le savoir humain ouvert et librement disponible pour l'humanité"

La technologie wiki

- 1990 : concept de *co-authoring* (C. M. Neuwirth)
- 1995 : **WikiWikiWeb** (W. Cunningham), impermanence [Langlais, 2014] :
lieu emblématique de "l'âge de la négociation" : il encourage une continuelle confrontation des opinions communes de ses utilisateurs

Naissance et premiers pas du projet Wikipédia – WP

Naissance du projet WP anglais

- **Jimmy Wales** : @Jimbo, finance, *trader*, Société *Bomis* qui gère des sites web *érotiques* financés par de la pub et projette la construction d'une encyclopédie "libre" : *Nupédia*, financé par la pub
- **Larry Sanger** : docteur en philo. recruté par Wales pour le projet *Nupédia*
- ***Nupédia*** : encyclopédie participative à laquelle tout auteur (qui a au moins un doctorat) peut proposer un article qui ne sera publié qu'après avoir été approuvé par des spécialistes du domaine, évalué en détail, discuté, expertisé par l'ensemble des participants du projet, envoyé à deux correcteurs non spécialistes, réévalué.
- Lourdeur du processus d'édition → faible productivité
- Sanger, connaissant WikiWikiWeb, propose de mettre en place un wiki pour faciliter le processus
- **15 janvier 2001, WP est ouvert** pour constituer l'espace de *brouillon* pour *Nupedia* (les meilleurs articles émanant de WP étant ensuite soumis au processus d'édition de *Nupédia*)

Naissance et premiers pas du projet WP

Un très gros boum

Février 2001	600 articles sont créés
Mars 2001	1300
Mai 2001	3900
Janvier 2002	20 000
Sept. 2003	WP compte plus de 100 000 article <i>Nupedia</i> est abandonné (seuls 24 articles ont été publiés)
2014	en France, WP est classée comme le 5e site le plus visité (21 millions de visiteurs par mois, 800 millions de pages visitées par mois) derrière <i>Google</i> , <i>Facebook</i> , <i>Youtube</i> , <i>Microsoft</i> et devant <i>Orange</i> , <i>Leboncoin</i> , <i>Amazon</i>

Premiers projets Wikipédia non anglais

- EN système de modération par un processus de consensus pur et de "dictateur bienveillant" (Wales, sorte d'éditeur encyclopédique)
- Wales : "les discussions sur la nature de Wikipédia ne peuvent figurer qu'en dehors de WP (de préférence sur la liste de diff.)"

WP français puis allemand et italien, un autre fonctionnement

Mise en place d'espaces de discussion et de processus de négociation (le dictateur bienveillant ne parle qu'anglais!) [Langlais, 2014]

- FR octobre 2002, création d'un page "prise de décision" par Florence Dévouard (@Anthère) où "le choix final dépend d'un vote consensuel et non d'un simple consensus"
- IT *Sondaggio*, "mécanisme simple et rapide pour résoudre facilement les problèmes actuels."
- DE *Meinungsbilder*, "clarifier les questions sur lesquelles aucun consensus n'a pu être atteint."

Un énorme objet d'étude

<http://wikipapers.referata.com>

"a compilation of resources [...] focused on the research of wikis"



Main page
Publications
Keywords
Authors
Datasets
Tools
Examples
...More lists...
Random page

Create new...
Publication
Author
Event
Keyword
Dataset
Journal
Tool

Activity
Community portal
Recent changes
New pages
RSS feeds
Follow us on
Twitter!

Page Discussion

Read

Edit

View history

Search

Q

List of publications

See also: *List of authors*, *List of datasets*, *List of tools*.

This is a **list of publications** available in WikiPapers. Currently, there are 6241 publications.

Filter by type:

- List of books (26) and List of book chapters (43)
- List of conference papers (4033)
- List of journal articles (1542)
- List of literature reviews (73)
- List of bachelor's theses (10), diploma theses (1), doctoral theses (52), master's theses (26)
- List of essays (11)
- List of peer-reviewed publications (661) and List of non peer-reviewed publications (17)
- List of magazine articles (29)
- List of unpublished works (4)

Filter by year:

- 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014

Filter by language:

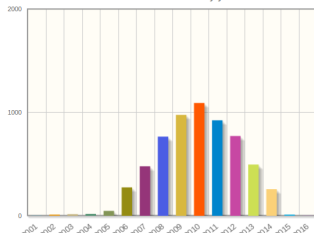
- Arabic, Catalan, Chinese, Dutch, English, French, Galician, German, Greek, Hungarian, Italian, Japanese, Polish, Portuguese, Russian, Slovenian, Spanish, Turkish

Filter by conference:

- CLEF, MathWikis, WikiAI, WikiSym, WikiViz

Create account Log in Log in / create account with OpenID

Publications distribution by year



(More trends stats#)

Un objet sociologique : "La négociation contre la démocratie" [Langlais, 2014]

"Démocratie consensuelle vs. démocratie majoritaire" [Langlais, 2014]

Ce que WP n'est pas : "WP n'est pas une expérience démocratique ni ne prétend s'apparenter à aucun autre système politique. La prise de décision et la résolution des conflits passent par deux processus fondamentaux (quoique non exclusifs) : l'édition des articles et l'exercice d'une discussion consensuelle — et non pas le vote."

Un observatoire de la collaboration entre humains

Wales, cité par [Langlais, 2014]

L'une des propriétés les plus surprenantes de WP c'est que l'on parvient à d'excellents résultats sans recourir à des règles formelles ou à des restrictions. Il y a des injonctions et des pressions sociales qui parviennent à garantir parfaitement le bon ordre des choses.

Une ressource pour le TAL

Exploitation des articles

- extraction de connaissances
[Zesch et al., 2008, Medelyan et al., 2009]
- construction de ressources multilingues [Kilgarrieff et al., 2010]

Exploitation des révisions

Un accès aux processus d'écriture [Ferschke et al., 2013]

Exploitation des discussions

Un accès aux processus de négociations
[Ferschke et al., 2012, Ferschke et al., 2013]

Exploitation des révisions et des discussions pour le TAL

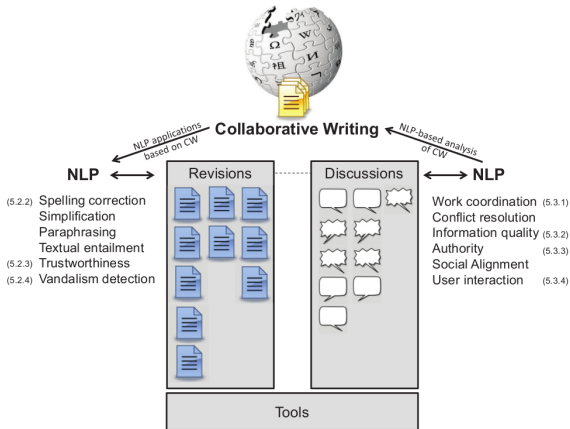


Fig. 5.1 The role of NLP in collaborative writing (CW): Topics covered in this chapter

Exploitation des révisions et des discussions pour le TAL

[Ferschke et al., 2013]

Exploitation des révisions

- *diff* entre les versions pour extraire des paires de variantes orthographiques, les paraphrases, les simplifications/résumés
- persistance d'une phrase comme indice d'importance
- détection du vandalisme (environ 7% des révisions dans la WP anglaise selon [Potthast et al., 2008])

Exploitation des discussions

Détection des conflits et désaccords (sujets à controverses)

WP As Corpus for CMC

Des genres et situations de communication plutôt bien définies

- articles encyclopédiques
- discussions
 - autour de la rédaction collaborative d'un article
 - autour du projet WP («cafés et bistrots», e.g. «Forum des Nouveaux», «Le salon de médiation», etc.)
- Journaux d'activité («Bulletin des patrouilleurs»)
- Ateliers et "Machines à café"

Accessibilité et quantité des données

- contenu libre distribué publiquement (Creative Commons by-sa)
- depuis 2001
- existe dans presque toutes les langues → objet d'étude international

Plan

- 1 Wikipédia comme objet d'étude
- 2 Profilage des discussions WP
 - Les discussions WP
 - Objectifs de l'étude et méthode
 - Le corpus WikiTalk
- 3 Classification of "Conflicting" vs. "Neutral" Discussions
- 4 A Bottom-Up Approach to Talk Page Profiling
- 5 Conclusion

La négociation à la base de l'écriture collaborative

Les pages de discussion WP, un besoin originel

Quelques jours après la création de WP, un des premiers contributeurs soulève le problème suivant : que faire des discussions concernant les articles ? Faut-il les déplacer dans une page à part ? Un autre contributeur pense que ce n'est pas nécessaire : sur un wiki tout se corrige de soi-même, les discussions trop étendues finiront par s'auto-réguler.

Larry Sanger ne tient pas le même discours : Je ne suis pas sûr que se référer à l'histoire du wiki soit très pertinent ici. Après tout, nous nous efforçons de créer une encyclopédie et, dès lors, nos articles ne peuvent pas être biaisés

[Langlais, 2014]

Consensus, (dés)accords et conflits dans la WP

Des conflits entre spécialistes et *wiki-anarchistes*

Les premiers mois d'existence de l'encyclopédie sont marqués par l'apparition de conflits entre contributeurs. Selon Sanger, des universitaires, venus du projet Nupedia, écrivant dans leur domaine de spécialité pour WP, auraient abandonné l'encyclopédie [...] lassés de croiser le fer avec des "wiki-anarchistes" d'un niveau d'expertise bien moindre. [Sahut, 2015, 241]

Des conflits entre auto-promotion et encyclopédie [Sahut, 2015, 264]

- [DiStaso, 2012] : 31% de professionnels de la communication ont contribué à l'écriture de la page de leur produit
- Wikiscanner, (V. Griffith) pour identifier les origines des adresses IP → Pepsi, CIA, Exxon, partis politiques dévoilés

Les pages de discussion WP

Un espace de négociation inédit

- Forum de discussion systématiquement associé à chaque article où les Wikipédiens peuvent discuter du processus d'écriture de l'article
- "The user discussions on the article Talk pages might shed light on this issue and give an insight into the otherwise hidden processes of collaboration that, until now, could only be analyzed via interviews or group observations in experimental settings." [Ferschke et al., 2012].

Un espace ouvert aux conflits, désaccords et consensus

- Antagonistic edits of the article structure and content
- Disagreements which may turn to conflicts when the editing and discussion processes are deadlocked [Poudat et al., 2016]

Exemple : <https://fr.wikipedia.org/wiki/Psychanalyse>

Les discussions WP as Corpus

Accessibilité et quantité des données

- (sorte de) Forum de discussion libre distribué publiquement (Creative Commons by-sa)
- existe dans une multitude de langues → objet d'étude international

Richesse des métadonnées

- thématique (portail thématique, article associé)
- connaissances partagées (article associé)
- indice de subjectivité (appel au calme, etc.)
- informations sur le locuteur (statut dans la communauté, participation à la WP, possibilité de profilage)

Objectifs

D'un point de vue global

- Description linguistique du genre "discussion en ligne" (CMR – Communications Médiées par les Réseaux) en partant des WikiDiscussions (relativement bien écrites)[Ho-Dac and Laippala, 2015]
- Définition de traits efficaces pour la classification des CMR
- Analyse linguistique de la négociation

Première étape présentée ici

- Identification de traits pertinents pour la sélection de pages propices aux négociations houleuses (à fort potentiel conflictuel)

Méthodologie

Testée

- ❶ Construction du WikiTalk Corpus en tirant un maximum partie des méta-données disponibles
- ❷ Description linguistique du WikiTalk corpus par des techniques du TAL :
 - classification automatique
 - fouille de données et profilage guidé par les données (analyse multidimensionnelle)

En gestation

- ❸ Définition des traits pertinents pour le profilage
- ❹ Sélection de discussions/fils propices aux négociations houleuses
- ❺ Annotation du caractère conflictuel des discussions/fils sélectionnées (consensus, désaccord, conflit)
- ❻ Apprentissage automatique pour la caractérisation de fils

Constitution du corpus WikiTalk

- ① Extraction des discussions depuis le dump
sauvegarde globale des pages courantes de la WP[FR] (archive
frwiki-20150512-pages-meta-current#.xml.bz2 diffusée librement sur la page
<http://dumps.wikimedia.org/frwiki/20150512/>)
- ② Sélection des discussions "à garder" selon le nombre de mots et de messages
- ③ Conversion selon la TEI-P5–CoMeRe
- ④ Analyse syntaxique automatique [Urieli, 2013]

3 487 480 /<title>Discussion/ dans le dump

User Talk pages	1,990,927	57%
Article Talk pages	1,496,553	43%
Redirections	116,432	8%
Empty talk pages (< 2 words)	1,013,791	68%
Selected talk pages	366,326	24%

Structuration et formatage du corpus

Structure de document encodée selon la TEI-P5

- fils – <div>
- titre de fil – <head>
- messages – <post who="user" when="timestamp" interactionLevel="#">

```

</p>
<p id="6"> Benjac 15 mai 2011 à 17:17 (CEST)</p>
</post>
<post id="2" who="Celette" bot="no" when="15-05-2011-17:16" interactionLevel="1">
  <p id="1">
    1) Aucun recul sur l'affaire, même pas une demi-journée.
  </p>
  <p id="2"> 2) "est une information condiscrable" : POV.</p>
  <p id="3">
    3) Vous avez une 4e contributrice pour vous barrer le chemin, moi. Les arguments sont clairs : aucun recul + l'introduction n'est pas un fourre-tout à polémique du moment. On attend de voir comment évolue l'affaire et on avise. Pour ma part, je suis prête à réverter.
  </p>
  <p id="4"> Cordialement, Celette 15 mai 2011 à 17:16 (CEST)</p>
</post>
<post id="3" who="Hatonjan" bot="no" when="15-05-2011-17:19" interactionLevel="2">
  <p id="1">(conflit d'edit)</p>
  <p id="2">
    Bonjour, avez vous lu RI ? C'est fort quand même vos remarque, à mon premier revert, mon commentaire vous renvoyait vers cette règle ! Et vous avez maintenu votre position malgré les reverts de 3 personnes que vous avez cité, sans rien apport dans votre obstination. Donc, déjà, ne nous accusez pas de discourtoisie. D'ailleurs le dialogue a été débuté par Lomita sur votre page de discussion.
  </p>
  <p id="3">
    Sur le fond, comme dit dans RI et Wikipédia:Conventions de styleBiographies qui sont les deux règles ici, c'est loin, très loin, d'être un fait majeur "pour le moment". Vous confondez une encyclopédie et un site de nouvelles. De plus, l'importance de cette information a déjà été débattu (et toujours en cours de débat) plus haut, et il a déjà été considéré que l'évènement ne justifie pas une partie en lui même au vu des informations à ce jour, alors dans le résumé, surement pas !
  </p>
  <p id="4">
    Bref, laissez la justice faire son travail, nous sommes peut être à l'aube d'un évènement majeur dans sa carrière, mais pour le moment, rien n'est fait. Hatonjan 15 mai 2011 à 17:19 (CEST)
  </p>
</post>
<post id="4" who="Lomita" bot="no" when="15-05-2011-17:24" interactionLevel="3">
  <p id="1">
    Entièrement d'accord avec mes collègues, je n'ai plus rien à rajouter, à part que ce fait divers est déjà traité dans l'article, à la place qui lui revient - --Lomita 15 mai 2011 à 17:24 (CEST)
  </p>

```

Aperçu global du corpus WikiTalk – bientôt disponible sous licence CC

talk pages	threads	posts	words
366,326	1,024,351	3,022,240	159,578,279

- 202,856 (55%) single post talks
- 181,503 (50%) under 53 words talks
- Some extremely long talks (up to 1,143 posts and 148,968 words)
- 150,603 (41%) monologues
- 40,413 (10%) talks involving between 8 and 228 different writers
- 80% anonymous posts
- Des auteurs très actifs

Aperçu global du corpus WikiTalk – bientôt disponible sous licence CC

talk pages	threads	posts	words
366,326	1,024,351	3,022,240	159,578,279

- 202,856 (55%) single post talks
- 181,503 (50%) under 53 words talks
- Some extremely long talks (up to 1,143 posts and 148,968 words)
- 150,603 (41%) monologues
- 40,413 (10%) talks involving between 8 and 228 different writers
- 80% anonymous posts
- Des auteurs très actifs

Des "activistes bienveillants"

en 2006, 90% de WP[FR] a été rédigé par 5% des participants

En 2015 0,5% (2250) – 3/4 de WP[FR] [Sahut, 2015, 263]

Contributions >=	Wikipedians		Contributions total	
1	491,679	100.0%	47,928,106	100.0%
3	206,703	42.0%	47,474,456	99.1%
10	107,596	21.9%	46,902,257	97.9%
32	45,699	9.3%	45,847,259	95.7%
100	19,581	4.0%	44,425,949	92.7%
316	9,139	1.9%	42,624,028	88.9%
1000	4,548	0.9%	40,070,097	83.6%
3162	2,250	0.5%	36,019,855	75.2%
10000	938	0.2%	28,560,465	59.6%
31623	256	0.1%	16,961,531	35.4%
100000	34	0.0%	5,449,567	11.4%
316228	2	0.0%	766,983	1.6%

contributions aux articles de WP[FR]

<https://stats.wikimedia.org/FR/TablesWikipediaFR.htm>

Méta-données exploitables (1)

Méta-données disponibles pour toute WikiDiscussion

- "discipline" i.e. portail thématique associé
- "avancement" i.e. qualité de l'article (évaluation par les Wikipédiens)
- "conflictiness" i.e. information saisie manuellement par les Wikipédiens via le *template* `{{keep calm}}` qui insère une bannière en haut de la page (existe également au niveau des articles, à creuser)



Please be [calm](#) and [civil](#) when you make comments or when you present evidence, and avoid [personal attacks](#). Please be patient as we work toward resolution of any issues in a peaceful, respectful [manner](#).

- "talk type" : information unique à la WP[FR] qui permet pour chaque article d'avoir plusieurs type de WikiDiscussion associé selon que l'on veuille aborder des questions de neutralité, qualité, etc.

Méta-données exploitables (1)

Méta-données disponibles à l'extérieur des WikiDiscussions

- méta-données extraites des articles
 - catégorie du portail thématique *Histoire*, *Art*, *Sport*, etc. : 11 grandes catégories (un même article peut être associé jusqu'à 7 catégories!)
 - bannières (actuellement non exploitées)

Affaire Bogdanoff



Cet article ne cite pas suffisamment ses sources (avril 2012).

Si vous disposez d'ouvrages ou d'articles de référence ou si vous connaissez des sites web de qualité traitant du thème abordé ici, merci de compléter l'article en donnant les **références utiles à sa vérifiabilité** et en les liant à la section « Notes et références » (modifier l'article, comment ajouter mes sources ?).

- l'existence d'une WikiDiscussions "parallèle" autour de problèmes de neutralité, qualité de l'article, etc. (liée à "talk types")

Autres discussions [\[liste\]](#)

Suppression - [Neutralité](#) - Droit d'auteur - Article de qualité - Bon article - Lumière sur - À faire - Archives



Autres discussions [\[liste\]](#)

Suppression - Neutralité - Droit d'auteur - [Article de qualité](#) - Bon article - Lumière sur - À faire - Archives



teiHeader actuel

```

- <teiHeader>
- <fileDesc>
  - <titleStmt>
    <title>3283822.xml</title>
  </titleStmt>
+ <publicationStmt></publicationStmt>
- <sourceDesc>
  - <biblStruct>
    - <analytic>
      <author>Wikipedia</author>
      <title>Discussion:Israël Shahak</title>
      <idno type="Wikipedia">3283822</idno>
    </analytic>
    + <monogr></monogr>
  </biblStruct>
</sourceDesc>
</fileDesc>
- <encodingDesc>
  <projectDecl/>
  - <classDecl>
    - <taxonomy>
      <bibl>Wikipedia</bibl>
      - <category type="genre">
        <catDesc>discussion article</catDesc>
      </category>
      - <category type="discipline">
        <catDesc>Seconde Guerre mondiale</catDesc>
        <catDesc>Israël</catDesc>
      </category>
      - <category type="avancement">
        <catDesc>BD</catDesc>
      </category>
      - <category type="interaction">
        <catDesc>{ Appel au calme }</catDesc>
      </category>
    </taxonomy>
  </classDecl>
</encodingDesc>
- <profileDesc>

```

Premiers points de sondage linguistiques

indice du niveau d'écriture [Baldwin et al., 2013]

- taux de mots inconnus (noms communs, verbes, adjectifs et adverbes que Talismane n'a pas réussi à lemmatiser)
- longueur moyenne des phrases et des mots

indice du degré de subjectivité

Affect : noms, adjectifs et verbes inclus dans le lexique des affects [Augustyn et al., 2008]

Pro1 : pronoms de première personne (*nous, je, j', me, m'*) en position sujet ou objet

Corpus de comparaison

Constitués

Corpus écrits	No tokens	Description
Forum Santé	236 368 151	Forum de discussion 2 585 188 messages
WikiDiscussion (2015)	132 406 816	Forum de discussion 3 022 240 messages
WikiArticles (2013)	226 207 672	Articles encyclopédiques

Parsés

Tous les corpus ont été analysés syntaxiquement avec Talismane (version 1.8.5b, beam de 5, modèle svn)

Premiers points de sondage linguistiques

	#Mots	mots inconnus	longueur moyenne		Affect	Pro1
		(%)	phrases	mots	%	%
Forum santé	236 368 151	22**	10,3	5,2	4,7	3,48
WikiDisc.	161 833 298	5*	18,2	5,4	2,1	0,05
Articles WP	226 207 672	5	14,6	5,5	1,8	1,34
	622 154 102	12	13,4	5,4	2,8	1,53

- * sur les 560 841 formes inconnues repérées 19% se retrouvent également dans les articles WP (i.e. ne sont pas nécessairement liés à des erreurs d'orthographe)
- ** 11,3/10 000 *jai* et *j* (comptés dans les inconnus) dans les forums de santé contre 0,2 dans les discussions
- *** à fortement nuancer, e.g. affect les plus fréquents pour chaque sous-corpus : *attendre* et *beau* dans les forums de santé, *croire* et *demande* dans les WikiDiscussions, *jouer* et *considérer* dans les articles WP.

[Ho-Dac and Laippala, 2015]

Formules d'ouverture des WikiDiscussions

Méthode : Extraction des n-grams ($n < 4$) de début de message

	Init msg/phr	%of Msg
Discussions	99.9	18
Avis	99.7	17
supprimer	97.1	6
conserver	95.7	5
Neutre	98.9	4
Votes	99.8	3
Signalé_par	98.8	3
des_articles_admissibles	99.2	3
Si_vous_êtes	15.6	2
Il_me_semble	32.8	2
Bilan	97.9	2
Il_y_a	23.4	2
Merci	28.2	2
Ce_n'_est	22.8	2
pourBA	98.7	2
Bon	53.6	2
Je_viens_de	60.8	2
Je_ne_vois	32.3	1
Je_suis_d'accord	59.7	1
En_effet	23.0	1
Je_pense_qu'	30.1	1
Effectivement	57.4	1
pour	88.1	1
Je_ne_sais	29.9	1

- Init msg/phr : propension des N-grams à apparaître en début de messages (p.r. en initiale de phrase)
- %of Msg : % de messages commençant pas ce N-gram

Formules d'ouverture des WikiDiscussions

Méthode : Extraction des n-grams ($n < 4$) de début de message

	Init msg/phr	%of Msg
Discussions	99.9	18
Avis	99.7	17
supprimer	97.1	6
conserver	95.7	5
Neutre	98.9	4
Votes	99.8	3
Signalé_par	98.8	3
des_articles_admissibles	99.2	3
Si_vous_êtes	15.6	2
Il_me_semble	32.8	2
Bilan	97.9	2
Il_y_a	23.4	2
Merci	28.2	2
Ce_n'_est	22.8	2
pourBA	98.7	2
Bon	53.6	2
Je_viens_de	60.8	2
Je_ne_vois	32.3	1
Je_suis_d'accord	59.7	1
En_effet	23.0	1
Je_pense_qu'	30.1	1
Effectivement	57.4	1
pour	88.1	1
Je_ne_sais	29.9	1

- Init msg/phr : propension des N-grams à apparaître en début de messages (p.r. en initiale de phrase)
- %of Msg : % de messages commençant pas ce N-gram
- si Init msg/phr > 85
⇒ **"code interne Wiki"**
- si Init msg/phr < 85
⇒ Argumentation et Recherche d'un consensus

Plan

- 1 Wikipédia comme objet d'étude
- 2 Profilage des discussions WP
- 3 Classification of "Conflicting" vs. "Neutral" Discussions**
- 4 A Bottom-Up Approach to Talk Page Profiling
- 5 Conclusion

Data-driven analysis of conflicts in WikiTalks : automatic classification

Using a classifier for both the automatic detection of conflicts and their analysis

- Linear classifier VowPal Wabbit [Agarwal et al., 2011]
- Supervised Machine Learning : Based on a set of *training data*, the classifier learns to detect the predefined classes and describe their most typical features
- Automatic classification already tested for classifying WP articles vs. talk pages ($F1 = 0.94$) and talk pages vs. eHealth Fora ($F1 = 1$)

Need for training data with a conflict class

- No available annotated corpus
- Available meta-data that may help for predefining a "conflict class"

An experimental classification of "conflicting" vs. "neutral" discussions

Training data : a distinction based on available meta-data

Conflict discussions	Neutral discussions
More than 100 words in the talk page	
From the header talk page : <i>Keep calm</i>	<i>A-class</i>
Existence of a parallel talk page about : article neutrality	article A-ranking
"Talk Type" : neutrality talk	na
2,028 pages 11 M words	4,569 pages 8.8 M words

Classification task : categorizing talk pages, threads or posts ?

Three available levels of granularity

- entire talk page : 6,597 units
- threads : 46,690 units
- posts : 194,289 units

But

- meta-data only available on the talk page level
- conflict seems to be more a matter of threads than of entire talk pages

Adhoc assumption

If a talk page is *a priori* categorized as conflicting, then the threads or posts it contains are also categorized as conflicting

Classification features : bag-of-lemmas or syntactic information ?

Two options

- A standard option : Lexical information [Scott et al., 2006]
- Problem : words reflect more the text topic than the text genre
- A more complex option : Syntactic information [Laippala et al., 2015]

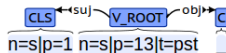


FIGURE – A delexicalized syntactic bi-arc describing a clitic+verb+conjunction as in the clause ‘I find that’.

Results (training on 2/3, test on 1/3)

features	threads			posts		
	P	R	F1	P	R	F1
lemmas	0.84	0.60	0.72	0.79	0.69	0.74
bi-arcs	0.55	0.48	0.52	0.63	0.59	0.61
units	46,690			194,289		

TABLE – Comparison of lexical vs. syntactic approaches for the automatic classification of conflict threads and posts.

- Syntactic approach lost
- (Adhoc) "conflict" posts are better classified than threads with the both approaches
- F1 is relatively low (79% of the detected posts are correctly classified, 69% are detected)

Lexical specificities of the conflicts

- Analysis of 40 most significant lemmas
- Two kinds : words referring to the writing process and those referring to article topics
 - style, to hope, respect, version, way of writing, restructuring, reformulation, neutralisation, clumsy, uncoherent, respect, mistake, controversy, debate, ok
 - rwanda, dictatorship, mandarin, quebec, islam, buddhism

Questioning

- the features used for classification
- our *a priori* classification of conflict and our Adhoc assumption

Plan

- 1 Wikipédia comme objet d'étude
- 2 Profilage des discussions WP
- 3 Classification of "Conflicting" vs. "Neutral" Discussions
- 4 A Bottom-Up Approach to Talk Page Profiling**
- 5 Conclusion

Profiling talk pages

- Exploring Rich Linguistic Features usually associated with interactional and rhetorical structures
- Mining talk pages and threads for discovering classes (without *a priori*)
- Identifying relevant classes for conflict analysis

Data mining method

Data mining tool

- R package FactoMineR dedicated to multivariate exploratory data analysis
- Principal Components Analysis (PCA) on talk pages and threads

Rich Linguistic Features

- **Global** : general quantitative characteristics of texts (talk pages and threads) e.g. number of words, number of contributors, presence of a "keep calm" banner ;
- **Thema** : portal sections of the associated article
- **Interact** : the frequency per texts of a wide range of interaction and politeness cues e.g. social deixis, marks of (dis)agreement, etc.
- **DiscRel** : the frequency per texts of connectives for each discourse relations as defined in the LEXCONN[Roze et al., 2012].

Global features

Information extracted from the talk page itself

logNnMots	number of words
nbFils	number of threads*
nbPosts	number of posts
profMax	"interactional depth"
nbContributeurs	number of different participants
nbAnonymes	number of anonymous posts
X.anonymes	% of anonymous posts
nbBots	number of posts written by bots
X.bots	% of posts written by bots
AdQ	"1" if the talk page is linked to a A-class article
polemique	"1" if the talk page has the banner "keep calm"

Thema features

WP section of the associated article

- 11 WP sections : art, geography, history, leisure, medicine, politics, religion, sciences, society, sport, technology
- Some articles are simultaneously in 7 sections !
- *Geography* is the most frequent section (119,359 talk pages)
- 11 features binarized (e.g. geography = 1/0)
- The same feature for talk pages and threads

Interact features

11 features automatically identified with simple regular expressions

Politeness	<i>thanks, hello, goodbye, hi, sincerely, cheers, please, would you, etc.</i>
Agreement	<i>OK, agree, yes, no, actually, etc.</i>
Question	<i>?</i>
Je	1st singular person pronouns + <i>personally</i>
Tu	2nd sing. pers. pronouns, informal "you"
Vous	2nd plur. pers. and formal "you" pronouns
Nous	1st plur. pers. pronouns
On	Informal "We"
WP	<i>Wikipedia</i> or <i>WP</i>
pour	Sentence-initial <i>For</i> or <i>I'm for</i>
contre	Sentence-initial <i>Against</i> or <i>I'm against</i>

DiscRel features

22 discourse relations with the number of identified connectives as value

- discourse relations as defined in the LEXCONN[Roze et al., 2012] : "a French lexicon of 328 discourse connectives, collected with their syntactic categories and the discourse relations they convey"
- When a connective is polysemious, all possible relations are considered
- alternation ; background ; commentary ; concession ; condition ; consequence ; continuation ; contrast ; detachment ; elaboration ; evidence ; explanation ; flashback ; goal ; narration ; opposition ; parallel ; rephrasing ; result ; summary ; temporality ; unknown relation

ACP parameters

- Considering only discussions with more than 100 words
- Only Interact and DiscRel features are taken into account (normalized on the number of words)
- The other features are just indicated (in blue) for permitting a global overview

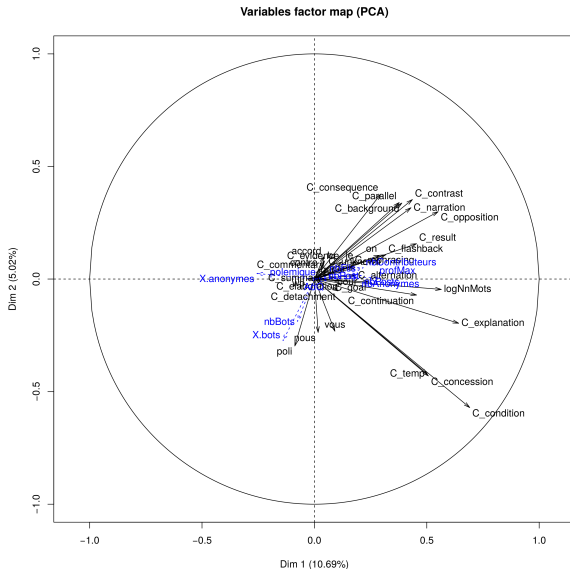
Results : ACP on linguistic features

- 5 dimensions that explain around 30% of the total variance
- A first dimension simply related to the number of words (the more words, the more features)
- A second dimension that differs acc. to the unit taken into account (thread or talk page)
- For **threads**, it opposes :
 - Dimension 2+ : more *I*, informal *we* (*on*) and discourse relations expressing **contrast**
 - Dimension 2- : agreement cues, formal *you* and discourse relations expressing alternation, consequence, goal and temporal relations

Results : ACP on linguistic features

- 5 dimensions that explain around 30% of the total variance
- A first dimension simply related to the number of words (the more words, the more features)
- A second dimension that differs acc. to the unit taken into account (thread or talk page)
- For **talk pages**, it opposes :
 - Dimension 2+ : more discourse relations expressing **contrast**, background/narration and causality
 - Dimension 2- : politeness cues, formal you and we and discourse relations expressing concession, condition and temporal relations

Results : ACP on linguistic features for talk pages



- Dimension 1 : the more words the more features
- Dimension 2+ : discourse relations expressing contrast, background/narration and causality
- Dimension 2- : politeness cues, formal you and we and discourse relations expressing concession, condition and temporal relations

Difficulties to go from these results to examples we may interpret

"few politeness cues, formal you and more discourse relations expressing contrast"

- Few politeness cues because few words or no real Interaction (one post per threads)
- Potentially only one formal you (perhaps included in a specific locution as "s'il vous plait" *please*)
- 17 connectives associated with contrast in the LexConn including the two very polysemous "but" and "while"

Example :

https://fr.wikipedia.org/wiki/Discussion:William_Tyndale

Plan

- 1 Wikipédia comme objet d'étude
- 2 Profilage des discussions WP
- 3 Classification of "Conflicting" vs. "Neutral" Discussions
- 4 A Bottom-Up Approach to Talk Page Profiling
- 5 Conclusion

Conclusion

- CMC genres are complex objects that challenge our traditional methods
- CMC genres require different levels of investigation
- Data mining techniques may give us some leads but...

Qualitative analyses and manual annotation are crucial

- A first annotation on two talk pages shows that only 50% of threads in two *a priori* conflict talk pages are conflict threads
- We must improve features that describe the thread level as for example by looking at the headings, the first post of the section and the context (ex : <https://fr.wikipedia.org/wiki/Discussion:Psychanalyse/arch1#choqu.C3.A9>)
- A lot of work in progress....



Agarwal, A., Chappelle, O., Dudik, M., and Langford, J. (2011).

A reliable effective terascale linear learning system.
JMLR, 15 :1111–1133.



Augustyn, M., Ben Hamou, S., Bloquet, G., Goossens, V., Loiseau, M., and Rynck, F. (2008).

Autour Des Langues Et Du Langage : Perspective Pluridisciplinaire, chapter Constitution de ressources pédagogiques numériques : le lexique des affects, page 407–414.
Grenoble : Presses Universitaires de Grenoble.



Baldwin, T., Cook, P., Lui, M., MacKinlay, A., and Wang, L. (2013).

How noisy social media text, how diffrent social media sources ?
In *IJCNLP*, pages 356–364.



DiStaso, M. W. (2012).

Measuring public relations wikipedia engagement : How bright is the rule.
Public Relations Journal, 6(2) :1–22.



Fersckhe, O., Daxenberger, J., and Gurevych, I. (2013).

A survey of nlp methods and resources for analyzing the collaborative writing process in Wikipedia.
In *The People's Web Meets NLP : Collaboratively Constructed Language Resources*. Springer.



Fersckhe, O., Gurevych, I., and Chebotar, Y. (2012).

Behind the article : Recognizing dialog acts in wikipedia talk pages.
In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 777–786. Association for Computational Linguistics.



Ho-Dac, L.-M. and Laippala, V. (2015).

Les discussions wikipedia : un corpus pour caractériser le genre "discussion".
In *International Research Days Social Media and CMC Corpora for the eHumanities*, Rennes, France.



Kilgarriff, A., Reddy, S., Pomikálek, J., and Avinesh, P. (2010).

A corpus factory for many languages.

In *LREC*.



Laippala, V., Kanerva, J., and Ginter, F. (2015).

Syntactic ngrams as keystuctures reflecting typical syntactic patterns of corpora in finnish.

Procedia - Social and Behavioral Sciences, 198 :233 – 241.



Langlais, P.-C. (2014).

La négociation contre la démocratie : le cas wikipedia.

Négociations, (1) :21–34.



Medelyan, O., Milne, D., Legg, C., and Witten, I. H. (2009).

Mining meaning from wikipedia.

International Journal of Human-Computer Studies, 67(9) :716–754.



Potthast, M., Stein, B., and Gerling, R. (2008).

Automatic vandalism detection in wikipedia.

In *Advances in Information Retrieval*, pages 663–668. Springer.



Poudat, C., Vanni, L., and Grabar, N. (2016).

How to explore conflicts in french wikipedia talk pages ?

In *JADT*, pages 645–656.



Roze, C., Danlos, L., and Muller, P. (2012).

Lexconn : A french lexicon of discourse connectives.

Discours, 10.



Sahut, G. (2015).

Wikipédia, une encyclopédie collaborative en quête de crédibilité : le référencement en questions.

PhD thesis, Université Toulouse Jean Jaurès ; Université de Toulouse.



Scott, M., , and Tribble, C. (2006).

Textual Patterns : Key Words and Corpus Analysis in Language Education.
Philadelphia, PA, USA : John Benjamins Publishing Company.



Urieli, A. (2013).

Analyse syntaxique robuste du français : concilier methods syntaxiques et connaissances linguistiques dans l'outil Talismane.

PhD thesis, Université de Toulouse – Jean Jaurès.



Zesch, T., Müller, C., and Gurevych, I. (2008).

Extracting lexical semantic knowledge from wikipedia and wiktionary.
In *LREC*, volume 8, pages 1646–1652.