

# Le Wiktionnaire : Construire un lexique contributif sur Internet

Regards croisés sur la lexicopoïèse

17 – 18 novembre 2017

Lucas Lévêque

 @Lyoko6

 Lyokoï Kun



Wiktionnaire  
*Le dictionnaire libre*

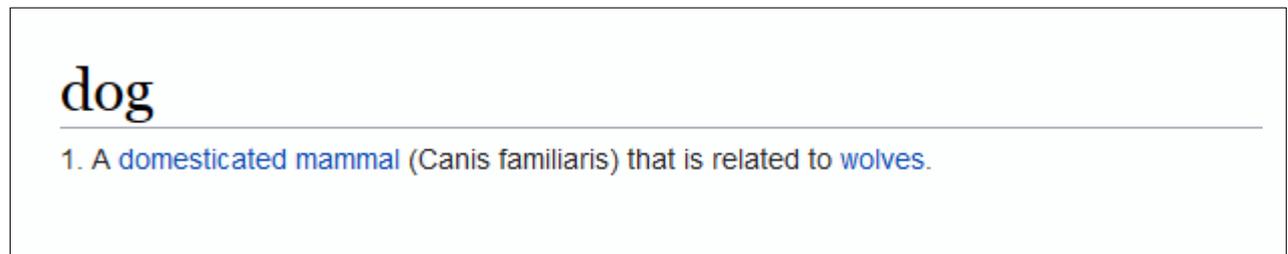
# Histoire du développement du Wiktionnaire

1. La création
2. 2006 : Premier tremplin
3. 2008 : Consolidations
4. 2012 : Changement d'échelle
5. Depuis 2012 : Passage à la vitesse supérieure



# 1. Le début : Wikipédia

- Initialement, un complément lexical sur Wikipédia
- Mais en contradiction avec la règle « Wikipédia n'est pas un dictionnaire »
- Les anglophones se lancent le 12 décembre 2002 sur un sous-domaine de Wikipédia
- Le 26 décembre 2002 le contributeur Daniel Alston, « Fonzy », déplace le projet sur son propre domaine informatique : [www.wiktionary.org](http://www.wiktionary.org).



# 1. Des discussions dès 2003 pour fr

- Un an et demi de discussion.
- Très peu de personnes s'impliquent, la discussion est courte.
- La première structure proposée d'un article, par Luna, a de grande ressemblance avec l'actuelle.
- Les logiques de fonctionnement sont déjà en place, la graphie en titre de page.
- On demande déjà un compte commun entre Wikipédia et le futur Wiktionnaire (Il arrivera en 2008).

# 1. En 2004, fr naît !

- Une décision commune est prise pour déplacer le Wiktionary anglophone vers en.wiktionary.org.
- Il est alors possible de créer d'autres versions linguistiques du projet.
- Le 22 mars 2004 sont créées la version française et la version polonaise.

Encyclopédie libre  
multilingue.

**Wiktionnaire**

/wik.sjɔ.nɛʁ/ *m.*  
Dictionnaire libre  
et universel.

Wikipedja

cyklopedia działająca w techno-  
logii → wiki

**Wikisłownik**

*m.* D. -a, N ~iem, blm:  
wolny słownik działa-  
jący w technologii wiki

wiklin|a - ż. C. ~ie, blm, bot.

# 1. Les débuts

- On est un peu chaotique :

1. accueil	21. bande dessinée	41. nouveau	61. mars	81. hypothèse
2. lire	22. oiseau	42. pinyin	62. avril	82. collaboration
3. encyclopédie	23. bird	43. sinogramme	63. mai	83. samedi
4. manga	24. allemand	44. kanji	64. juin	84. dimanche
5. ouvrage	25. Allemande	45. chair à canon	65. juillet	85. prendre
6. siège	26. Allemands	46. Friday	66. octobre	86. le
7. chaise	27. jeudi	47. Freitag	67. novembre	87. la
8. fauteuil	28. vendredi	48. Montag	68. décembre	88. les
9. meuble	29. ailé	49. jour	69. computer	89. 星期
10. armchair	30. wing	50. Dienstag	70. new	90. fin
11. mardi	31. voler	51. Mittwoch	71. year	91. abréviation
12. lundi	32. fly	52. Donnerstag	72. anglais	92. essai
13. semaine	33. steal	53. groupe	73. lieu	93. Monday
14. militaire	34. ovipare	54. école	74. physique	94. take
15. suis	35. poisson	55. CD-ROM	75. concertation	95. March
16. barbe à papa	36. fish	56. photographie	76. interrogation	96. Thursday
17. manchot	37. travail	57. ordinateur	77. homogène	97. June
18. pingouin	38. armée	58. janvier	78. imprimante	98. Tuesday
19. penguin	39. work	59. année	79. empirique	99. 火
20. mercredi	40. army	60. février	80. observation	100. 人

# 1. Dès l'ouverture, un usage pédagogique

## Enoncé [\[URL\]](#) [\[lien\]](#)

---

### Partie 1 [\[URL\]](#) [\[lien\]](#) ↑

---

Trouvez les définitions de deux termes parmi ceux qui sont écrits en rouge. Placez ces définitions dans un document sur votre ordinateur. Si des mots ne sont pas clairs et explicites pour vous, reformulez avec vos propres mots.

**format d'image**, collaboration, **jpg**, image vectorielle, ordinateur, photographie, école, imprimante, **cdrom**, disquette, plan, répartir, groupe, physique, chute, bille, rapport, rédiger, sciences, interrogation, empirique, observation, hypothèse, capture, homogène, concertation, adaptation, explicite reformuler

Pour les 6èmes **variable**, cellule, macro, constante

### Partie 2 [\[URL\]](#) [\[lien\]](#) ↑

---

- Expliquez comment vous avez trouvé ces définitions
- Comment vous êtes sûr qu'elles sont correctes ? (autrement dit: pourquoi leur faites-vous confiance)

### Partie 3 [\[URL\]](#) [\[lien\]](#) ↑

---

- Revenez sur cette page (ou rafraichissez/actualisez la)
- Regardez si un des mots que vous aviez choisi est toujours en rouge, cliquez dessus et entrez votre définition.
- Vérifiez votre orthographe.
- dans le bas, vous verrez une ligne "résumé". écrivez "premier essai de ...votre nom..."
- Et cliquez sur "enregistrer"
- Observez ce qui se passe.
- Essayez de découvrir le site. De comprendre "ce que c'est" et "à quoi ça sert".
- Si vous voulez entrer d'autres définitions, allez d'abord vous inscrire en cliquant sur le lien identification en haut à gauche.

# 1. La progression est fulgurante

- 100 articles le 26 mars 2004
- 200 articles le 1er avril 2004
- 300 articles le 13 avril 2004
- 2000 articles le 6 janvier 2005
- 5000 articles le 27 avril 2005
- 10 000 articles le 28 août 2005
- 15 000 articles le 2 octobre 2005
- 22 000 articles le 29 octobre 2005

## 2. Bouleversement en 2006

- La plupart des fondateurs et fondatrices, et des contributeurs et contributrices de la première heure n'y sont plus.
- Dès 2006, ce sont d'autres qui prennent la main et plusieurs de ceux-ci sont encore des piliers aujourd'hui.
- Le DAF8, libre de droit, est importé au début de l'année (fin de l'import 21 février 2006). La vérification des pages concernées sera finie 4 ans après.
- Du 3 janvier au 29 septembre, le Wiktionnaire francophone est le premier Wiktionnaire en nombre de pages de contenu.

Date	Wikipédiens				Articles				Base de données			Liens				
	total	nouveau	rédactions		nombre	nouveaux par jour	moyenne		rédactions	taille	mots	interne	interwiki	image	externe	redirections
			> 5	> 100			rédactions	octets								
jan 2006	224	17	59	15	128 K	838	2,5	266	73 K	47 Mo	4,3 M	355 K	103 K	2,5 K	2,1 K	977
déc 2005	207	15	43	7	102 K	2044	2,4	290	103 K	39 Mo	3,7 M	442 K	58 K	399	198	838

## 2. La technique démarre

- On se charge de modéliser les prononciations pour qu'elle soit lisible par des machines.
- On lance la catégorisation automatique des mots. Le gain de temps est considérable !
- On facilite la gestion du plurilinguisme des pages.
- On conventionne le projet d'une base solide.
- En août 2006, la barre des 200 000 pages est franchie, un an avant c'était celle des 10 000.
- Grâce à l'hébergement des fichiers multimédias sur Commons, le Wiktionnaire gagne 2 100 illustrations (35 000 aujourd'hui).

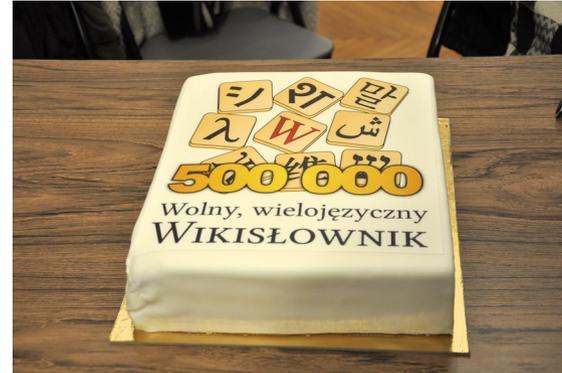
## 2. Une communauté qui mute

- On double le nombre de contributeurs et contributrices très actives (+100 modifications par mois).
- 17 contributeurs et contributrices arrivent, l'équipe s'agrandit naturellement et beaucoup sont des passionné.e.s du sujet.
- Les premières rencontres ont lieu, deux fois, à Paris. Beaucoup de sujets y sont discutés et cela facilite grandement les échanges entre contributeurs et contributrices.



# 3. 2008 : Consolidations

- Nouvelle réunion
  - On discute des rimes
  - Des annexes
  - De autant/au temps pour moi
  - Des gros ajouts au projet
- 1 millions d'articles en octobre
- Réorganisation des langues et des définitions
- Le Wiktionnaire est très proche de sa forme actuelle.

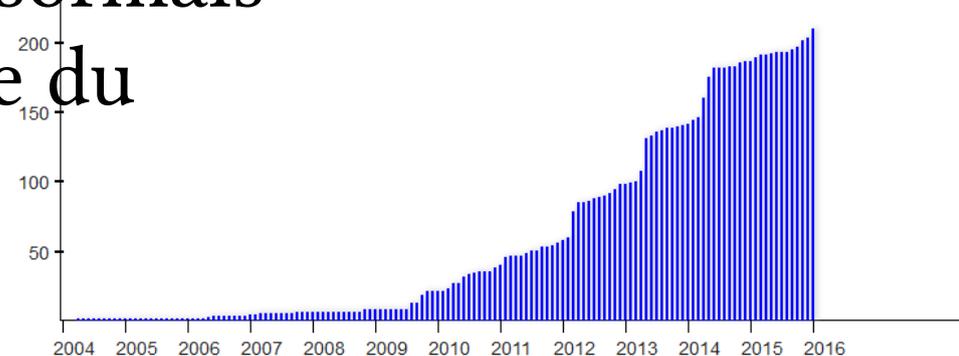


## 4. 2012 : on change d'échelle

- Les codes de langues sont massivement ajoutés, on passe les 2500 langues.
- On passe à la version moderne de la structure des pages.
- Le Wiktionnaire évolue désormais sur le qualitatif en parallèle du quantitatif.

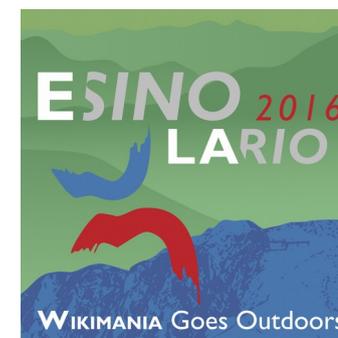


Wiktionnaire  
*Le dictionnaire libre*



# 5. Ces 4 dernières années, l'accélération

- 2013 : Nouvelle réunion
  - On discute de l'intégration du langage LUA
  - Les nouveaux piliers se rencontrent pour la première fois
  - On se découvre et on reprend contact
- 2013 : Shtooka déverse 100 000 fichiers audio de mots
- 2014 : Première demande de conférence
- 2014 : premiers ateliers
- 2015 : première permanence mensuelle au KoToPo
- 2016 : Conférence à la Wikimania



# 5. Où en est-on aujourd'hui ?

En 2017, les chiffres explosent :

- 3 100 000 pages
- 360 000 lemmes en français
- 530 000 définitions en français
- 1 600 000 formes avec flexions
- 4 110 langues
- 22 langues avec +10 000 définitions
- 20 conférences ont été données
- 30 ateliers de formations
- 30 000 audios de prononciations
- 32 800 médias d'illustration

- Formalisation de la première convention avec un institut de recherche : Logoscope.
- Un mensuel d'actualité
- Un projet de plus en plus reconnu au sein du mouvement et présent sur les réseaux sociaux.
- Une équipe importante de contributeurs et contributrices
- Un savoir et un savoir-faire important et affirmé
- Des projets internes qui structurent les différents champs du savoir linguistique

L o g  s c o p e 2

# 5. Où irons-nous demain ?

- La pertinence de disposer d'un développeur dédié au développement technique du projet est de plus en plus discutée. Il y a beaucoup de chose à faire.
- L'amélioration de l'accueil des nouveaux est une ligne majeur de plusieurs réflexions qui implique l'intégration de nouvelles technologies et savoir-faire (tutoriel vidéo, MOOC, didacticiel interactif...).
- Évidemment, le projet est toujours demandeur de nouveaux outils pour contribuer et/ou faire découvrir son contenu.
- La venue de LUA et d'autres solutions technique apportent une nouvelle vision de l'automatisation : Traitement automatisé des rimes, génération améliorée des flexions pour les langues, organisation des articles...
- Intercommunication de plus en plus forte avec les autres projets  
Wikimédia : Wikisource, Wikimédia commons, Lingua Libre, Wikipédia, Wikiversité, Wikilivres....

# Merci de votre attention.

