

La place des langues de France hexagonales dans les projets Wikimedia

Rapport de Wikimedia France –
association pour le libre partage de la connaissance

pour

la Délégation générale à la langue française
et aux langues de France

Juillet 2014



Table des matières

1 Introduction.....	3
2 Présentation des projets Wikimedia.....	3
2.1 Wikipédia.....	3
2.2 Wiktionnaire.....	4
2.3 Wikisource.....	5
2.4 Wikiquote.....	6
2.5 Wikibooks.....	6
2.6 Wikinews.....	7
3 Politique linguistique des projets Wikimedia.....	7
3.1 Ouverture de versions linguistiques et comité des langues.....	7
3.2 Problématique de la définition et des limites des langues.....	8
4 Panorama des langues de France sur les projets Wikimedia.....	8
4.1 Aperçu global des langues de France.....	8
4.2 Traduction de l’interface.....	10
4.3 Discussion par langue.....	11
4.3.1 Catalan.....	11
4.3.2 Basque.....	12
4.3.3 Occitan.....	13
4.3.4 Breton.....	13
4.3.5 Luxembourgeois.....	14
4.3.6 Alémanique.....	15
4.3.7 Wallon.....	15
4.3.8 Corse.....	16
4.3.9 Flamand occidental.....	16
4.3.10 Normand.....	17
4.3.11 Picard.....	18
4.3.12 Francoprovençal.....	18
5 Conclusion.....	19

1 . Introduction

Le présent rapport a pour objet la description et une tentative d'analyse de la place qu'occupent les langues de France hexagonales au sein des projets Wikimedia. Il est rédigé par des Wikimédiens versés dans les langues (sans toutefois être linguistes professionnels).

Les projets Wikimedia sont apparus sur Internet en 2001 avec la création de son principal projet, Wikipédia. Ils sont hébergés par une fondation américaine à but non lucratif, s'enrichissent de manière bénévole et se financent presque intégralement par les dons. Ces projets sont multilingues et comptent à cet égard près de 300 langues en 2014, parmi lesquelles 12 langues de France hexagonales. Les langues qui seront au cœur de ce rapport sont les langues présentes sur les projets Wikimedia : le français (pour comparaison avec les autres langues de France), le catalan, le basque, l'occitan, le breton, le luxembourgeois, l'alsacien, le wallon, le corse, le flamand occidental, le normand, le picard et le francoprovençal.

La première partie du rapport présentera les projets Wikimedia et leur politique linguistique, et la seconde partie reprendra de manière globale puis détaillée la place de chaque langue de France au sein des projets concernés.

2 . Présentation des projets Wikimedia

Les projets Wikimedia sont un ensemble de 11 projets de collecte et de dissémination de la connaissance basés sur le principe du wiki. Ils sont hébergés par une fondation américaine à but non-lucratif, la Wikimedia Foundation. Outre leur caractéristique commune d'être créés de façon collaborative par les internautes, ils sont également disponibles sous une licence libre permettant de réutiliser les contenus très facilement au regard du droit d'auteur, y compris pour des réutilisations commerciales.

Ci-après sont présentés les 6 projets Wikimedia pour lesquels il existe au moins une langue régionale de France. Ce sont, pour les projets ayant des versions linguistiques indépendantes : Wikipédia, Wiktionnaire, Wikisource, Wikiquote, Wikibooks et Wikinews. Outre ces projets, Wikivoyage (guide de voyage) et la Wikiversité (cours en ligne) n'ont pas de version linguistique en langue de France autre que le français, et trois projets sont multilingues : Wikimedia Commons (médiathèque contenant actuellement 21 millions d'images et autres documents, dont l'interface peut être traduite en n'importe quelle langue), Wikidata (base de connaissances sémantiques contenant actuellement 15 millions d'entités, pouvant être décrites dans n'importe quelle langue) et Wikispecies (inventaire du vivant : les noms vernaculaires des espèces peuvent être indiqués dans n'importe quelle langue).

2.1 Wikipédia

Wikipédia est un projet d'encyclopédie universelle, multilingue (287 langues mi-2014) et sous licence CC-BY-SA, créée par Jimmy Wales et Larry Sanger le 15 janvier 2001. Cette encyclopédie a pour principale caractéristique d'être écrite de façon collaborative par tous les internautes qui désirent y contribuer. À sa création, cinq principes fondateurs ont été définis pour toutes les langues :

- Wikipédia est une encyclopédie : elle recense les sujets notables et n'a pas vocation à publier n'importe quel sujet ;
- il est recherché la neutralité de point de vue : cela passe par la citation des sources et l'exposition de tous les points de vue notables ;
- Wikipédia est sous licence libre : réutilisation beaucoup plus souple que le droit d'auteur classique ;
- respect des règles de savoir-vivre : les agressions sont à proscrire et le consensus doit être de mise ;
- il n'y a pas d'autres règles : en anglais *be bold* (soyez audacieux), il ne faut pas avoir peur de faire des erreurs.



WIKIPÉDIA
L'encyclopédie libre

Chaque version linguistique est indépendante des autres, ce qui résulte en des histoires différentes et en des règles de fonctionnement interne différentes. Bien que tout le monde puisse contribuer sur tous les articles, on distingue plusieurs types de contribution au projet :

- écriture d'articles, ajout de contenu à partir d'autres sources reconnues, ajout d'illustrations, discussions autour du contenu et de sa présentation,
- maintenance des articles : mise en forme des articles, réécriture selon le style encyclopédique, classement des articles dans des catégories, curation du contenu sur des portails thématiques, suppression des articles trop peu notables, suppression des violations des textes violant le droit d'auteur, etc.
- surveillance des modifications afin de prévenir toute dégradation du contenu, cela se fait en observant les modifications récentes, avec l'aide de programmes automatiques et semi-automatiques,
- organisation de la vie communautaire : accueil des nouveaux contributeurs, rédaction des pages d'aide, création et maintenance des règles internes liées au contenu ou à la vie communautaire, résolution des conflits, etc.

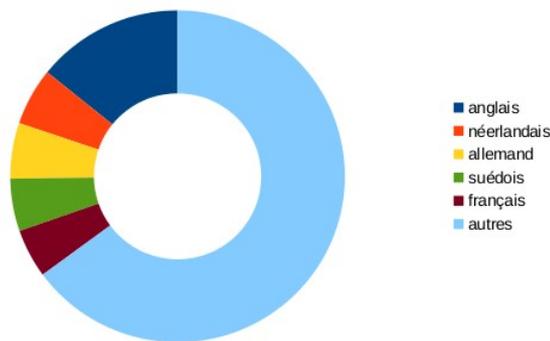


Illustration 1: Taille relative des Wikipédias en nombre d'articles :

- anglais : 14 % ;
 - néerlandais, allemand, suédois, français : ~ 5 % ;
 - autres : 65 %

Le site wikipedia.org (toutes langues confondues) se place régulièrement autour de la 5^e place en fréquentation, dans le monde (selon les statistiques d'Alexa.com) et en France (selon les statistiques de Mediametrie.fr). La Wikimedia Foundation donne le chiffre de 80 000 contributeurs réguliers par mois toutes langues et projets confondus d'après ses statistiques (au moins 5 contributions par mois),¹ et la Wikipédia en français compte entre 700 et 5 000 contributeurs mensuels réguliers selon le niveau d'activité spécifié (de 5 à plus de 100 modifications par mois).²

La création d'articles de façon automatisée ou semi-automatisée fait l'objet de débats au sein de la communauté Wikipédia. Récemment, par exemple, la Wikipédia en suédois a créé de façon automatisée un grand nombre d'articles, la faisant entrer dans le top 5 en nombre d'articles. Les promoteurs de la création automatisée soutiennent que cela permet de créer facilement des contenus qui n'existeraient pas autrement, et les opposants disent que cela crée une charge d'entretien et de surveillance pour la communauté, en particulier si celle-ci est trop petite. Une position intermédiaire est de présenter des « articles virtuels », n'existant pas en tant que tel, mais présentant des données de base sur le sujet grâce à Wikidata, lorsque le sujet existe sur ce site.

2.2 Wiktionnaire

Le Wiktionnaire est un projet de dictionnaire libre et gratuit, que chacun peut construire et qui décrit les mots, locutions, sigles, préfixes, suffixes et expressions de toutes les langues ainsi que les symboles et les caractères (appelés simplement mots pour simplifier). Il y a 171 versions linguistiques du Wiktionnaire ; il y a eu des propositions visant à rassembler tous les Wiktionnaires en un seul site multilingue (ayant mené en particulier à la création du site Omegawiki.org en 2004, indépendant des projets Wikimedia) mais le statu quo d'une multitude de versions linguistiques reste aujourd'hui



1 http://reportcard.wmflabs.org/graphs/active_editors

2 Statistiques générales sur Wikipédia en français, <http://stats.wikimedia.org/FR/TablesWikipediaFR.htm>

pour les Wiktionnaires.³

Chaque version linguistique possède des définitions rédigées dans la langue correspondante, mais les entrées peuvent aussi comprendre des mots d'autres langues. Par exemple, au 12 juin 2014, la version francophone du Wiktionnaire compte plus de 2 500 000 articles, parmi lesquels plus de 1 300 000 mots en français, langue qui occupe la première place. Les premières langues de France qui sont représentées sur le Wiktionnaire francophone sont l'occitan avec plus de 17 000 mots (20^e place) et le catalan avec plus de 15 000 mots (21^e place).⁴ Cela est également vrai pour les versions en langues régionales, qui peuvent également compter des mots en français ou d'autres langues. Il existe des Wiktionnaires en langue catalane (58 000 mots, 37^e place), en langue basque (47 000 mots, 40^e place), en langue bretonne (33 000 mots, 45^e place), en langue occitane (24 000 mots, 52^e place), en langue wallonne (12 000 mots, 71^e place), en langue corse (2 000 mots, 98^e place) et en langue alémanique (0 mot).

L'article d'un mot comporte une section pour chacune des langues dans lesquelles il existe. Chacune de ces sections contient :

- la prononciation ;
- l'étymologie, les différents sens, des exemples et citations ;
- des antonymes et des synonymes, entre autres ;
- des conjugaisons et déclinaisons ;
- des traductions vers les autres langues ;
- des mots appartenant à un domaine commun.

De plus, chacun peut également, sans même devoir se créer un compte :

- ajouter des mots manquants, des traductions, des étymologies (sans enfreindre le droit d'auteur) ;
- corriger facilement les erreurs éventuelles ;
- discuter sur une définition ou une prononciation ;
- participer à la vie communautaire ;
- proposer des améliorations.

2.3 Wikisource

Wikisource est un projet de bibliothèque publié sous licence libre auquel chacun peut contribuer. Cette bibliothèque compte des contenus d'ouvrages passés dans le domaine public ou publiés sous licence libre, et son principal travail est de vérifier la reconnaissance optique de caractères (OCR) pour donner des livres plein texte facilement réutilisables. Wikisource contient 63 versions linguistiques et une version multilingue ; les versions linguistiques s'occupent des livres principalement écrits dans leur langue et la version multilingue contient des livres multilingues, les livres écrits dans une langue n'ayant pas de Wikisource dans la langue, et elle sert d'incubateur avant d'ouvrir une Wikisource indépendante dans une langue donnée.⁵



Trois principes généraux délimitent le contenu de Wikisource :

- des textes déjà publiés ;
Les documents placés sur Wikisource doivent avoir été déjà publiés ailleurs, sur papier.
Deux exceptions au critère de l'édition sur papier : les publications uniquement électroniques ne sont autorisées que si elles proviennent de sites officiels ou d'éditeurs en ligne reconnus (c'est-à-dire qui ont une obligation de dépôt légal auprès de la bibliothèque nationale de leur pays) et les traductions.

³ Cette présentation est en partie issue de la page d'accueil du Wiktionnaire en français, sous licence CC-BY-SA 3.0. https://fr.wiktionary.org/w/index.php?title=Wiktionnaire:Page_d%E2%80%99accueil&oldid=15173062

⁴ Statistiques du Wiktionnaire en français, <https://fr.wiktionary.org/wiki/Wiktionnaire:Statistiques>

⁵ Cette présentation est en partie issue de la page « Wikisource:Qu'est-ce que Wikisource ? » sur la Wikisource en français, sous licence CC-BY-SA 3.0. https://fr.wikisource.org/w/index.php?title=Wikisource:Qu%E2%80%99est-ce_que_Wikisource_%3F&oldid=4457940

- des textes dans le domaine public ou sous licence libre ;
Les documents placés sur Wikisource doivent être dans le domaine public ou sous une licence libre compatible avec la licence Creative Commons CC-BY-SA (qui remplace la licence GFDL anciennement utilisée).
- des textes conformes à une édition ; Les documents placés sur Wikisource doivent être objectivement identiques à une édition de référence.

Ce qui est autorisé sur Wikisource :

- livres, manuscrits, presse et revues, chansons et livrets ;
- textes originaux libres de droits et déjà publiés auparavant. Le texte de référence publié est de préférence la dernière version intégrale revue par l’auteur, mais plusieurs versions significatives d’un même texte peuvent cohabiter ;
- traductions libres de droits ou réalisées par des contributeurs, à condition que ces derniers ne s’approprient pas la traduction, qui est une œuvre collective de Wikisource ;
- données statistiques telles que les résultats d’élections ;
- bibliographies des auteurs présents dans Wikisource.

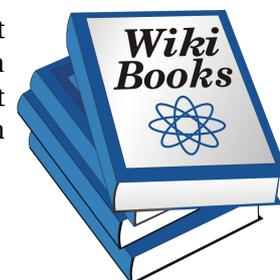
2.4 Wikiquote

Wikiquote est un projet multilingue et collaboratif de recueil de citations en ligne publié sous licence libre. Wikiquote a pour but d’établir une base de données de référence en termes de citations. Wikiquote se décline en 91 versions linguistiques. En nombre de pages (chaque page contenant les citations d’une personne en particulier), la plus grosse Wikiquote est celle en polonais (24 000 pages), celle en français contient 6 300 pages (8^e place), et les autres langues de France présentes sont le catalan (1 700 pages, 21^e place), le basque (220 pages, 50^e place), le breton (73 pages, 60^e place), le corse (6 pages), l’alémanique (1 page).



2.5 Wikibooks

Wikibooks (parfois appelé « Wikilivres » en français) est un projet lancé le 10 juillet 2003 sous le nom de Wikimedia Free Textbook Project et destiné à créer un ensemble d’ouvrages pédagogiques librement accessibles via un wiki. Wikibooks est une collection de textes pédagogiques libres rassemblés en livres, écrits en collaboration.



Ce que Wikibooks contient :⁶

- des livres pédagogiques structurés ;
Contrairement à Wikipédia, Wikibooks permet de créer un livre complet sur un sujet, et non une simple page synthétique. Les livres complètent donc Wikipédia en proposant un contenu plus complet, structuré, uniforme et plus orienté vers l’apprentissage. Le contenu des Wikibooks est en effet structuré en différentes parties ou chapitres, distribuées sur plusieurs pages généralement, mais toujours organisées de manière à guider le lecteur. Cela démarque les livres des articles encyclopédiques qui sont censés être autonomes et centrés sur un sujet précis.
- des livres d’apprentissage ;
Wikibooks contient des livres d’apprentissage dans tous les domaines. Il existe des livres pour apprendre différentes langues, des livres d’enseignement pour tous les arts (musique, dessin, photographie...), des livres sur l’informatique etc.
- des livres scolaires ;
Les livres pédagogiques contiennent bien entendu des livres scolaires, qui pourront aider les élèves, les étudiants à appréhender les enseignements qu’ils suivent.

6 Cette présentation est en partie issue de la page « Wikilivres:Présentation » de Wikibooks en français, sous licence CC-BY-SA 3.0.

<https://fr.wikibooks.org/w/index.php?title=Wikilivres:Pr%C3%A9sentation&oldid=432966>

- des textes de savoir-faire ;
Le savoir-faire est un savoir appréciable qu’il est important de transmettre. Qu’il s’agisse de cuisine, de formation techniques, de programmation ou de bricolage, ce genre de livre a tout à fait sa place sur Wikibooks.
- des textes annotés ;
Il est possible d’annoter des textes littéraires (si ceux-ci sont dans le domaine public).

Wikibooks se décline en 121 versions linguistiques. La plus importante est l’anglais avec près de 50 000 pages, le français se place en troisième position avec 16 000 pages, et les langues de France comptent le catalan (3 200 pages, 15^e place), l’occitan (118 pages, 62^e place), le basque (94 pages, 65^e place), le corse (12 pages), l’alémanique (8 pages) et le wallon (3 pages).

2.6 Wikinews

Wikinews est un projet de sources d’information et d’actualité auquel chacun peut contribuer. Créé en 2004, son but est d’être une source d’informations journalistiques sous licence libre. Outre le français, la seule langue de France présente sur ce projet est le catalan.



3 . Politique linguistique des projets Wikimedia

3.1 Ouverture de versions linguistiques et comité des langues

L’ouverture, et parfois la fermeture, de nouvelles versions linguistiques de projets Wikimedia (Wikipédia, Wiktionnaire, etc.) sont gérées par un comité de la Wikimedia Foundation, le comité des langues⁷ (language committee en anglais). Celui-ci est composé d’une quinzaine de Wikimédiens et linguistes qui évaluent les candidatures des nouvelles langues : définition claire de la langue, adéquation par rapport à quelques règles, viabilité du projet dans cette langue.

L’ouverture d’un projet dans une langue est conditionnée à l’existence d’un code ISO 639 (parties 1 à 3),⁸ autrement dit une langue doit être suffisamment délimitée par les linguistes et répertoriée par cette norme. Lorsqu’une langue donnée n’existe pas dans le registre de cette norme, un important travail en amont doit être accompli, d’une part pour que la langue soit suffisamment unifiée pour être candidate, donc qu’une communauté parlant cette langue et qu’un corpus écrit existent, et d’autre part que des linguistes et des communautés parlant cette langue aient soutenu l’inscription dans ce registre.

Parmi les quelques autres règles qui doivent être satisfaites pour l’ouverture d’un projet dans une langue donnée. il faut qu’il n’existe pas de langue plus générale (par exemple une requête pour l’ouverture d’une Wikipédia en Hochdeutsch serait très probablement refusée car cela se confond suffisamment avec l’allemand codifié par la langue ISO 639-1 de). Une autre règle dit qu’il n’est plus possible d’ouvrir des Wikipédias dans des langues mortes (cela ne s’applique pas forcément aux autres projets Wikimedia, par exemple des Wikisource en grec ancien ou en vieux norrois pourraient être ouvertes) ; la Wikipédia en latin, langue généralement considérée comme morte, a été ouverte en 2003 et fait exception à cette règle.

La viabilité d’un projet Wikimedia dans une langue donnée s’estime d’une part par le nombre de locuteurs référencés par les linguistes et d’autre part empiriquement par l’achèvement de deux travaux de base :

- la traduction de l’interface du logiciel, au moins pour les messages de base (environ 500 messages, généralement courts),
- une petite base d’articles réalisée sur un incubateur, permettant de voir le nombre de contributeurs, leur régularité, leur maîtrise de la langue, leurs difficultés ou facilités à créer des articles, etc.

⁷ Comité des langues, https://meta.wikimedia.org/wiki/Language_committee

⁸ Politique de proposition de langue, https://meta.wikimedia.org/wiki/Meta:Language_proposal_policy

3.2 Problématique de la définition et des limites des langues

Wikipédia – principalement mais également les autres projets Wikimedia – est aujourd’hui un des principaux exemples de projets multilingues,⁹ ou plus exactement de projets dans un grand nombre de langues : presque 300 actuellement sur un total d’environ 6 000 à 7 000 langues (Ethnologue.com recense 7 106 langues mi-2014). Partant de ce constat, certaines personnes dans le mouvement Wikimedia soutiennent l’idée d’une augmentation forte du nombre de langues couvertes par les projets Wikimedia, par exemple de multiplier par 10 le nombre de langues couvertes, passant à 3 000 langues.

Un tel développement se heurte à plusieurs problèmes :

- la viabilité des langues : il faut qu’un certain nombre de locuteurs parlent bien la langue, aient suffisamment de temps à consacrer à l’écriture du projet Wikimedia et aient un accès correct à l’Internet ;
- l’écriture des langues : actuellement, les seules langues présentes sur les projets Wikimedia sont écrites ; cela exclut de fait une certaine partie des langues orales ou visuelles comme les langues des signes, même s’il est possible que des projets en langues des signes voient le jour – il y a des propositions d’ouverture de tels projets ; d’autre part l’écriture des langues est facilitée lorsqu’elle est un minimum standardisée et qu’un corpus écrit existe ;
- la définition et la délimitation des langues : plusieurs langues, et en particulier en France métropolitaine et d’outre-mer, s’intègrent dans un continuum linguistique dont il est difficile de tracer des limites précises ;
- la codification ISO 639 des langues : comme mentionné précédemment, il est maintenant obligatoire pour une langue d’avoir un code ISO 639 1-3 pour candidater à l’ouverture d’un projet Wikimedia.

Pour les langues de France n’ayant pas de projets Wikimedia, ces problématiques peuvent également s’appliquer. Par exemple, le gallo n’a pas de code ISO 639, ce qui l’exclut de fait des projets Wikimedia. Pour plusieurs autres langues, il peut être difficile d’un point de vue linguistique d’affirmer que certains dialectes sont suffisamment délimités pour être qualifiés de langue et recevoir ainsi un code ISO 639. Pour les langues de France, cette problématique peut concerner par exemple le lorrain, le franc-comtois, le champenois, le mayennais, le poitevin-saintongeais, etc. L’un des rédacteurs de ce rapport a également fait en 2011 un petit état des lieux de la présence des langues de France vis-à-vis du projet Wikisource.¹⁰

Quand une langue possède un code ISO 639 pour candidater à l’ouverture d’un projet, une grande variété de dialectes différents peut rendre la gestion interne du projet assez délicate. Parmi les projets Wikimedia faisant cohabiter plusieurs dialectes, le normand et l’occitan peuvent être cités. La Wikipédia en langue normande indique par ailleurs sur chaque page le dialecte utilisé pour sa rédaction.

4 .Panorama des langues de France sur les projets Wikimedia

4.1 Aperçu global des langues de France

L’encyclopédie Wikipédia est le projet de la Wikimedia Foundation qui compte le plus de langues de France. Au nombre de 19, ces langues sont l’alémanique (inclut l’alsacien), le basque, le breton, le catalan, le corse, le créole haïtien, le francoprovençal (ou arpitan), le flamand occidental, le luxembourgeois, le normand, l’occitan, le picard, le tahitien et le wallon. S’ajoutent également pour les langues non territoriales les versions de Wikipédia en arabe dialectal (inclus dans la Wikipédia en arabe), en arménien occidental (inclus dans la

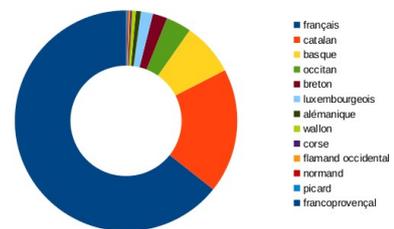


Illustration 2: Taille du français par rapport aux autres langues de France.

9 Mentionné par Daniel Prado dans l’émission Place de la toile du 3 mars 2012.

<http://www.franceculture.fr/emission-place-de-la-toile-les-langues-de-l-internet-2012-03-03>

10 Liste de langues de France vis-à-vis du projet Wikisource,

<https://fr.wikisource.org/wiki/Utilisateur:VIGNERON/langues>

Wikipédia en arménien), en judéo-espagnol, en romani et en yiddish.

Parmi ces langues, l'encyclopédie Wikipédia est donc rédigée en 12 langues de France hexagonales et occupe à ce titre la première place des projets Wikimedia.

Le second projet selon le nombre de langues est le Wiktionnaire, avec huit langues régionales que sont le basque, le breton, le catalan, le corse, le luxembourgeois, l'occitan, le picard et le wallon. Arrive ensuite Wikisource, avec cinq langues régionales que sont le basque, le breton, le catalan, le corse et l'occitan, et Wikiquote avec également cinq langues régionales que sont le basque, le breton, le catalan, le corse et le luxembourgeois. Enfin, Wikibooks compte trois langues que sont le basque, le catalan et l'occitan, et une version de Wikinews existe uniquement en catalan.

Le basque et le catalan sont donc les deux langues de France hexagonales les plus représentées au sein des projets Wikimedia. Quantitativement, c'est-à-dire en nombre de pages existant dans la totalité des projets, le catalan est en tête avec plus de 500 000 pages, suivi du basque avec plus de 200 000 pages, et de l'occitan avec plus de 100 000 pages, le breton comptant pour sa part un peu plus de 85 000 pages.¹¹ Le catalan compte, par ailleurs, à lui seul plus de la moitié des articles écrits dans toutes les langues de France sur Wikipédia (voir illustration 3).

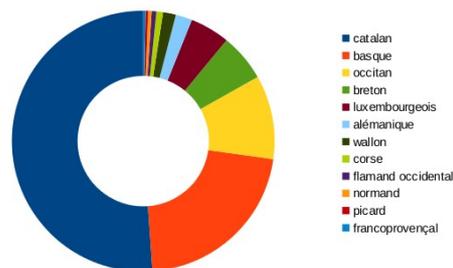


Illustration 3: Taille relative des langues de France.

Dans tous les cas, l'ensemble des langues de France hexagonales sont représentées en juin 2014 sur plus d'un million de pages de contenu (n'incluant par exemple pas les pages de discussion ou les pages techniques) sur l'ensemble des projets Wikimedia (sur un total général de 73,5 millions de pages de contenu, toutes langues et tous projets confondus).

Nombre de pages par langue et projet Wikimedia

Langue	Wiki-pédia	Wik-tionnaire	Wiki-source	Wiki-quote	Wiki-books	Wiki-news	Total
Alémanique	17 000 +	0	0	0	0	0	17 000
Basque	182 000 +	47 000 +	50	200 +	100 -	0	230 000
Breton	50 000 -	33 000 -	3 500 -	73	0	0	87 000
Catalan	430 000 +	55 000 +	19 000 +	1 700 +	3 200 +	4 400 +	515 000
Corse	6 700 +	2 100 +	1	6	12	0	8 800
Flamand occidental	5 000 -	0	0	0	0	0	5 000
Francoprovençal	2 300 +	0	0	0	0	0	2 300
Luxembourgeois	41 000 +	6 000 +	0	21	0	0	47 000
Normand	3 500 +	0	0	0	0	0	3 500
Occitan	90 000 -	25 000 -	100 +	0	120 -	0	115 000
Picard	2 700 +	250 -	0	0	0	0	2 900
Wallon	13 000 +	12 200 +	0	0	0	0	25 200
Total approximatif	845 000	181 000	23 000	2 000	3 500	4 400	1 060 000

Note : les chiffres inscrits en **gras italique** correspondent à des pages qui ne possèdent pas, ou plus, leur propre wiki.

11 Liste des projets Wikimedia par taille, https://meta.wikimedia.org/wiki/List_of_Wikimedia_projects_by_size

Le catalan, le basque et l'occitan étant également parlés dans d'autres pays (l'Espagne pour chaque ainsi que l'Italie pour l'occitan), la langue parlée exclusivement en France la plus présente sur les projets Wikimedia est donc le breton. Cette langue peut de toute manière être considérée comme l'unique langue de France dans ce cas, puisque toutes les autres langues présentes sur les projets Wikimedia se situent à proximité d'une frontière d'un pays dans lequel elles sont également parlées. Ceci concerne, en plus de l'Espagne, tous les pays de la Belgique à l'Italie ainsi que les îles anglo-normandes pour le normand.

Aussi, à l'exception de l'occitan qui couvre une grande partie du territoire sud du pays, une large zone s'étendant de la région Poitou-Charentes aux régions Lorraine et Franche-Comté ne présente actuellement aucune langue ni dialecte de France sur les projets Wikimedia. Cette zone ne compte quasiment que des langues d'oïl, avec par exemple l'angevin, le berrichon, le bourbonnais, le bourguignon-morvandiau, le champenois, le franc-comtois, le gallo, le lorrain, l'orléanais, le poitevin-saintongeais ou le tourangeot, mais s'y ajoute aussi le francique en Lorraine.

4.2 Traduction de l'interface

MediaWiki, le logiciel de wiki utilisé sur l'ensemble des autres projets Wikimedia et sur d'autres wikis indépendants, est actuellement disponible en 314 langues, ou plutôt codes de langues. Sur ces 314 codes de langues et outre les langues à part entière, certains codes notent des variétés de registre (`de-formal`, `nl-informal`), mais la plupart des autres exceptions dénotent des écritures de langues différentes, par exemple pour effectuer des translittérations automatiques entre les écritures chinoises ou entre les alphabets cyrillique, latin et arabe (serbe, kazakh, kurde, etc.). Une fois les doublons d'écriture retirés, on trouve que MediaWiki est traduit en 278 langues, avec toutefois des niveaux d'avancement différents.

La traduction entière de MediaWiki comprend actuellement 19 331 messages, parmi lesquels 3 046 pour le logiciel de base sans les extensions et 16 271 pour les diverses extensions. Bien sûr, les sites fonctionnant avec MediaWiki n'utilisent jamais toutes les extensions ; pour les sites Wikimedia, de part les extensions activées, il faut compter 8 859 messages à traduire. Pour traduire entièrement l'interface des sites Wikimedia, cela demande donc actuellement la traduction de 11 905 messages (3 046 + 8 859). Les traductions sont faites sur le site <https://translatewiki.net> (site indépendant de la Wikimedia Foundation), dont l'objectif initial de traduire MediaWiki s'est élargi à la traduction de projets et logiciels libres (aujourd'hui 27 projets y sont traduits).

Le tableau ci-dessous présente les statistiques de traduction pour les langues de France hexagonales. On peut distinguer trois groupes par rapport à l'achèvement des traductions :

1. forte traduction : le logiciel de base est traduit à 90 % et le total traduit à 50 %, cela concerne le français, le luxembourgeois, le breton, le catalan, l'occitan, le francoprovençal, l'alémanique et le basque ;
2. traduction moyenne : le logiciel de base est partiellement traduit, cela concerne le wallon, le picard et le corse ;
3. faible traduction : cela concerne le flamand occidental et le normand.

Outre l'utilisation de MediaWiki sur les projets Wikimedia, MediaWiki peut être utilisé pour tout autre site, et une bonne traduction du logiciel peut permettre à certaines communautés linguistiques (en langues de France ou non) de l'utiliser. L'investissement dans la traduction du logiciel ne concerne donc pas seulement Wikimedia, mais plus largement les communautés linguistiques. Translatewiki.net avait par exemple sponsorisé durant une semaine en 2008 une trentaine de traducteurs très actifs en partageant entre eux 1 000 euros¹² ; dans la même veine, Wikimedia Sverige a organisé il y a quelques mois un *translation sprint* rassemblant sept participants bénévoles pendant une après-midi pour améliorer la traduction de MediaWiki en suédois.¹³

12 Page de Translatewiki.net, https://translatewiki.net/wiki/Translating:Language_project/500claim

13 Article du blog de la Wikimedia Foundation, <http://blog.wikimedia.org/2013/11/08/wikimedia-sweden-translation-sprint/>

Langue	Manquants				Faits			
	Tout	Media-Wiki	Toutes ext.	Ext. WM	Tout	Media-Wiki	Toutes ext.	Ext. WM
Total					19 331	3 046	16 271	8 859
Français	0	0	0	0	100 %	100 %	100 %	100 %
Luxembourgeois	3 623	48	3 573	1 728	81 %	98 %	78 %	80 %
Breton	5 917	95	5 820	3 048	69 %	97 %	64 %	66 %
Catalan	7 689	5	7 682	3 308	60 %	100 %	53 %	63 %
Occitan	7 839	83	7 753	4 180	59 %	97 %	52 %	53 %
Francoprovençal	8 194	382	7 809	3 820	58 %	87 %	52 %	57 %
Alémanique	9 255	224	9 028	4 671	52 %	93 %	45 %	47 %
Basque	13 009	245	12 761	6 649	33 %	92 %	22 %	25 %
Wallon	16 826	1 552	15 271	7 926	13 %	49 %	6 %	11 %
Picard	18 520	2 303	16 214	8 804	4 %	24 %	0 %	1 %
Corse	18 607	2 468	16 135	8 730	4 %	19 %	1 %	1 %
Flamand occidental	19 190	2 927	16 256	8 844	1 %	4 %	0 %	0 %
Normand	19 227	2 957	16 259	8 847	1 %	3 %	0 %	0 %

4.3 Discussion par langue

Les langues sont listées ci-après par ordre d'importance selon le nombre d'articles sur Wikipédia, la langue la plus importante étant le catalan et la moins importante le francoprovençal.

4.3.1 Catalan

La version de Wikipédia en catalan¹⁴ (« Viquipèdia ») a été ouverte le 16 mars 2001, en utilisant le code *ca*. Il s'agit de la deuxième version linguistique de Wikipédia lancée par la Wikimedia Foundation, deux mois après le lancement de la version anglophone et une semaine avant celui de la version francophone.

Elle atteint en juin 2014 les 430 000 articles, rédigés par environ 500 contributeurs actifs (c'est-à-dire effectuant plus de cinq modifications par mois) sur près de 10 000 contributeurs inscrits, soit un équivalent d'environ 58 contributeurs pour un million de locuteurs. Le projet compte par ailleurs chaque mois environ une centaine de nouveaux contributeurs et de 60 à 80 nouveaux articles.



VIQUIPÈDIA
L'enciclopèdia lliure

La page d'accueil de Wikipédia en catalan est consultée en moyenne environ 10 000 fois par jour. Durant les trois premières semaines du mois de juin 2014, les dix articles les plus consultés, hors pics de consultation ponctuels non représentatifs, étaient « Coupe du monde de football de 2014 », « Nuit de la Saint-Jean », « Barcelone », « Catalogne », « Éducation physique », « Moyen Âge », « Espagne », « Seconde guerre mondiale », « Wikipédia en catalan » et « Guerre de succession espagnole ».

¹⁴ <https://ca.wikipedia.org>

Le version du Wiktionnaire en catalan¹⁵ (« Viccionari ») a été lancée en mai 2004. Le projet compte en juin 2014 plus de 55 000 définitions et 5 750 utilisateurs, dont 30 ont été actifs au cours des 30 derniers jours.

La version de Wikisource en catalan¹⁶ (« Viquitexts ») a été lancée en juin 2006. Le projet compte en juin 2014 plus de 19 000 pages, issues de près de 500 livres et auteurs. Parmi ces pages, 5 800 ont été validées par plusieurs personnes (29 %) et 13 600 ont été révisées par au moins une personne (69 %). Près de 3 400 contributeurs y sont enregistrés, dont 15 ont été actifs au cours des 30 derniers jours.

La version de Wikiquote en catalan¹⁷ (« Vikidites ») a été lancée en août 2004. Le projet compte plus de 1 700 pages d’auteurs contenant une ou plusieurs de leurs citations et 3 000 contributeurs, dont 17 ont été actifs au cours des 30 derniers jours.

La version de Wikibooks en catalan¹⁸ (« Vikilibres ») a été lancée en 2004. Le projet compte plus 3 200 pages de 155 livres en juin 2014 et près de 3 000 contributeurs, dont 12 ont été actifs au cours des 30 derniers jours.

La version de Wikinews en catalan¹⁹ (« Vikinotícies ») a été lancée en 2006. Elle compte 4 400 articles en juin 2014 et plus de 2 600 contributeurs dont 9 ont été actifs au cours des 30 derniers jours.

4.3.2 Basque

La version de Wikipédia en basque²⁰ (« Wikipedia ») a été ouverte en décembre 2001, en utilisant le code eu. Elle fait également des premières versions linguistiques de Wikipédia lancées en 2001.

Elle atteint en juin 2014 les 182 000 articles, rédigés par plus de 50 000 utilisateurs enregistrés, dont 232 ont été actifs au cours des 30 derniers jours, soit un équivalent d’environ 65 contributeurs pour un million de locuteurs.

La page d’accueil de Wikipédia en basque est consultée en moyenne environ 2 000 fois par jour. Durant les trois premières semaines du mois de juin 2014, les dix articles les plus consultés, hors pics de consultation ponctuels non représentatifs, étaient « Wiki », « Basque », « Wikipédia », « Pays basque », « Moyen Âge », « Espagne », « États-Unis d’Amérique », « Wikipédia en basque », « Union européenne » et « Abraham Lincoln ».

La version du Wiktionnaire en basque²¹ (« Wikiztegia ») a été lancée en 2006. Le projet compte plus de 47 000 pages en juin 2014 et 2 200 contributeurs inscrits, dont 11 ont été actifs durant les 30 derniers jours.

La version de Wikisource en basque²² (« Wikiturriak ») ne possède pas son propre nom de domaine, mais une sous-page du projet Wikisource, créée en 2007. Elle compte 50 pages en juin 2014.

La version de Wikiquote en basque²³ (« Wikiesanak ») a été lancée en 2004. Le projet compte plus de 200 pages de citations et proverbes en juin 2014.



WIKIPEDIA
Entziklopedia askea

15 <https://ca.wikipedia.org>

16 <https://ca.wikisource.org>

17 <https://ca.wikiquote.org>

18 <https://ca.wikibooks.org>

19 <https://ca.wikinews.org>

20 <https://eu.wikipedia.org>

21 <https://eu.wiktionary.org>

22 https://wikisource.org/wiki/Main_Page/Euskara

23 <https://eu.wikiquote.org>

La version de Wikibooks en basque²⁴ (« Wikiliburuak ») a été lancée en 2005. Elle compte près de 100 pages en juin 2014.

4.3.3 Occitan

La version de Wikipédia en occitan²⁵ (« Wikipèdia ») a été ouverte en décembre 2003, en utilisant le code oc. Elle compte en juin 2014 près de 90 000 pages, rédigées par plus de 20 000 contributeurs, dont 88 ont été actifs durant les 30 derniers jours, soit l'équivalent d'environ 12 contributeurs pour un million de locuteurs.



WIKIPÈDIA
L'enciclopèdia liura

Wikipédia en occitan inclut différents dialectes. La règle qui y est établie en matière de cohabitation de ces dialectes est la suivante : « Tout article de la Wikipédia en occitan doit être rédigé dans l'une des grandes variétés régionales de l'occitan, ainsi définies : auvergnat, gascon (incluant l'aranais), limousin, languedocien, provençal (incluant le niçois) et le vivaro-alpin. Toutes ces variétés régionales sont égales, et la Wikipédia en occitan valorise la diversité dialectale. »²⁶ À notre connaissance, il n'existe pas de moyen de connaître la répartition de chacun des dialectes permis sur la version en occitan de Wikipédia. Toutefois, une liste non-exhaustive présente sur le wiki indique que 28 Wikipédiens rédigent en languedocien, 23 en gascon et 14 en provençal, les autres dialectes étant rédigés par moins de 10 contributeurs revendiqués.²⁷

La page d'accueil de Wikipédia en occitan est consultée en moyenne environ 1 800 fois par jour. Durant les trois premières semaines du mois de juin 2014, les dix articles les plus consultés, hors pics de consultation ponctuels non représentatifs, étaient « Occitan », « Union européenne », « Wikipédia », « France », « Russie », « Royaume-Uni », « Turquie », « Italie », « Allemagne » et « Pologne ».

La version du Wiktionnaire en occitan²⁸ (« Wikiccionari ») a été lancée en 2004. Elle compte près de 25 000 pages en juin 2014 et plus de 2 000 contributeurs inscrits, dont 11 ont été actifs au cours des 30 derniers jours.

La version de Wikisource en occitan²⁹ (« Wikisource ») ne possède pas son propre nom de domaine, mais une sous-page du projet créée en 2005. Elle compte un peu plus de 100 pages en juin 2014.

La version de Wikibooks en occitan³⁰ (« Wikilibres ») a été lancée en 2006. Elle compte environ 120 pages en juin 2014 et un peu plus de 1 000 contributeurs inscrits, dont 2 ont été actifs au cours des 30 derniers jours.

4.3.4 Breton

La version de Wikipédia en breton³¹ (« Wikipedia ») a été ouverte en juin 2004, en utilisant le code br. Elle compte en juin 2014 près de 50 000 pages, rédigées par plus de 30 000 contributeurs dont 85 ont été actifs au cours des 30 derniers jours. Avec l'équivalent d'environ 92 contributeurs pour un million de locuteurs, cette version de Wikipédia est de ce point de vue la plus active.

²⁴ <https://eu.wikibooks.org>

²⁵ <https://oc.wikipedia.org>

²⁶ « Variétés régionales de l'occitan », https://oc.wikipedia.org/wiki/Wikipèdia:Carta_lingüistica#Varietats_regionalas_de_l.27occitan

²⁷ Catégorie:Wikipédien par dialecte, https://oc.wikipedia.org/wiki/Categoria:Wikipedian_per_dialècte

²⁸ <https://oc.wiktionary.org>

²⁹ https://wikisource.org/wiki/Main_Page:Occitan

³⁰ <https://oc.wikibooks.org>

³¹ <https://br.wikipedia.org>

La page d'accueil de Wikipédia en breton est consultée en moyenne environ 1 800 fois par jour. Durant les trois premières semaines du mois de juin 2014, les dix articles les plus consultés, hors pics de consultation ponctuels non représentatifs, étaient « Allemagne », « Wikipédia », « Japon », « Autriche », « Italie », « Belgique », « 2005 », « Brésil », « Breton » et « République populaire de Chine ».

La version du Wiktionnaire en breton³² (« Wikeriadur ») a été lancée en 2005. Elle compte près de 33 000 pages en juin 2014 et plus de 2 000 contributeurs inscrits, dont 10 ont été actifs au cours des 30 derniers jours.

La version de Wikisource en breton³³ (« Wikimammenn ») a été initiée en 2007 et possède son propre projet depuis novembre 2010. En juin 2014, elle compte près de 3 500 pages correspondant à 148 livres. Un millier de contributeurs sont inscrits, dont 5 ont été actifs au cours des 30 derniers jours. Ses contributeurs ont effectué plus de 65 000 modifications en un peu moins de 4 ans.

La version de Wikiquote en breton³⁴ (« Wikiarroud ») a été lancée en 2007. Elle compte 73 pages de citations et un millier de contributeurs inscrits, dont 2 ont été actifs au cours des 30 derniers jours.

4.3.5 Luxembourgeois

La version de Wikipédia en luxembourgeois³⁵ (« Wikipedia ») a été ouverte en juillet 2004, en utilisant le code lb. Elle compte en juin 2014 plus de 41 000 articles et plus de 500 contributeurs, dont une trentaine effectuent plus de cinq modifications par mois, soit l'équivalent d'environ 82 contributeurs pour un million de locuteurs (le deuxième niveau le plus élevé après la version de Wikipédia en breton).

La page d'accueil de Wikipédia en luxembourgeois est consultée en moyenne environ 1 000 fois par jour. Durant les trois premières semaines du mois de juin 2014, les dix articles les plus consultés, hors pics de consultation ponctuels non représentatifs, étaient « Luxembourg (pays) », « Luxembourgeois », « Bartreng », « Bertrange (Luxembourg) », « Gréiwelshaff », « Findelshaff », « États-Unis d'Amérique », « Wikipédia », « Allemagne », « Italie » et « Royaume-Uni ».

La version du Wiktionnaire en luxembourgeois³⁶ (« Wiktionnaire ») a été lancée en 2006. Elle compte plus de 6 000 pages en juin 2014 et plus de 1 500 contributeurs inscrits, dont 8 ont été actifs au cours des 30 derniers jours.

La version de Wikiquote en luxembourgeois³⁷ (« Wikiquote ») a été lancée en 2006. Elle compte 21 pages de citations en juin 2014. Près de 600 contributeurs y sont inscrits, mais aucun n'y a contribué au cours des 30 derniers jours.



WIKIPEDIA
An holloueziadur digor



WIKIPEDIA
Déi fräi Enzyklopedie

32 <https://br.wiktionary.org>

33 <https://br.wikisource.org>

34 <https://br.wikiquote.org>

35 <https://lb.wikipedia.org>

36 <https://lb.wiktionary.org>

37 <https://lb.wikiquote.org>

4.3.6 Alémanique

La version de Wikipédia en alémanique³⁸ (« Wikipedia ») a été ouverte en novembre 2003, en utilisant le code als qui correspondait à l'origine à l'alsacien, mais finalement élargie à l'alémanique en raison de sa trop lente progression. Les projets en alémanique utilisent le code als pour des raisons historiques bien que cela ne soit pas correct vis-à-vis de la norme ISO 639-3 (où als désigne le tosqe, dialecte de l'albanais). Le code ISO 639-3 correct est le code gsw, qui est toutefois présent dans les attributs lang du HTML, permettant une détection correcte de la langue par les outils informatiques.

Elle compte en juin 2014 plus de 17 000 articles et plus de 38 000 contributeurs inscrits, dont 81 ont été actifs au cours des 30 derniers jours, soit l'équivalent d'un contributeur pour un million d'habitants.

La page d'accueil de Wikipédia en alémanique est consultée, en moyenne, environ 50 fois par jour. Durant les trois premières semaines du mois de juin 2014, les dix articles les plus consultés, hors pics de consultation ponctuels non représentatifs, étaient « Europe », « Allemagne », « Suisse », « Sexualité », « États-Unis d'Amérique », « Alémanique », « France », « Espagne », « Anglais » et « Coupe du monde de football de 2014 ».

Des projets similaires au Wiktionnaire³⁹ (« Werterbuach »), à Wikisource⁴⁰ (« Textsammlong »), à Wikibooks⁴¹ (« Lehrbiacher ») et à Wikiquote⁴² (« Spruchsammlong ») ont été lancés directement sur le site de Wikipédia en alémanique. Ces projets ont pu exister sur des projets indépendants pendant une très courte période, avant que décision ne soit finalement prise de fusionner chacun de leur contenu sur Wikipédia en alémanique, en 2008.

4.3.7 Wallon

La version de Wikipédia en wallon⁴³ (« Wikipedia ») a été ouverte en juin 2003, en utilisant le code wa.

Elle compte en juin 2014 plus de 13 000 articles et près de 10 000 contributeurs inscrits, dont 13 ont été actifs au cours des 30 derniers jours, soit l'équivalent d'environ 3 contributeurs pour un million de locuteurs.

La page d'accueil de Wikipédia en wallon est consultée en moyenne environ 300 fois par jour. Durant les trois premières semaines du mois de juin 2014, les dix articles les plus consultés, hors pics de consultation ponctuels non représentatifs, étaient « Wallon », « Wikipédia », « États-Unis », « Pronom », « Adolf Hitler », « Anglais (langue) », « Allemagne », « Mer », « Belgique » et « France ».

La version du Wiktionnaire en wallon⁴⁴ (« Wiccioneaire ») a été lancée en 2005. Elle compte plus de 12 200 pages en juin 2014 et plus de 1 500 contributeurs, dont 11 ont été actifs au cours des 30 derniers jours.



WIKIPEDIA
Di frei Enzyklopedy



WIKIPEDIA
Li libe eciclopedeye

38 <https://als.wikipedia.org>

39 <https://als.wikipedia.org/wiki/Wort:Houptsyte>

40 <https://als.wikipedia.org/wiki/Text:Houptsyte>

41 <https://als.wikipedia.org/wiki/Buech:Houptsyte>

42 <https://als.wikipedia.org/wiki/Spruch:Houptsyte>

43 <https://wa.wikipedia.org>

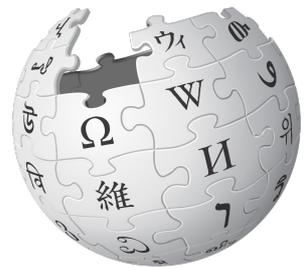
44 <https://wa.wiktionary.org>

4.3.8 Corse

La version de Wikipédia en corse⁴⁵ (« Wikipedia ») a été ouverte en novembre 2003, en utilisant le code **co**.

Elle compte en juin 2014 plus de 6 700 articles et plus de 8 600 contributeurs inscrits, dont 23 ont été actifs au cours des 30 derniers jours, soit l'équivalent d'environ 12 contributeurs pour un million de locuteurs.

La page d'accueil de Wikipédia en corse est consultée en moyenne environ 400 fois par jour. Durant les trois premières semaines du mois de juin 2014, les dix articles les plus consultés, hors pics de consultation ponctuels non représentatifs, étaient « Wikipédia », « France », « Langue corse », « Langue anglaise », « Italie », « Corse », « Espagne », « Russie », « États-Unis d'Amérique » et « Langue allemande ».



WIKIPEDIA
Enciclopedia libara

La version du Wiktionnaire en corse⁴⁶ (« Wikizziunariu ») a été lancée en décembre 2007. Elle compte plus de 2 100 pages en juin 2014 et 1 700 contributeurs, dont 3 ont été actifs au cours des 30 derniers jours.

La version de Wikisource en corse⁴⁷ (« Wikisource ») ne possède pas son propre nom de domaine, mais une sous-page du projet lancée en 2003. Elle compte en juin 2014 une seule page, la Déclaration universelle des droits de l'homme.⁴⁸

La version de Wikiquote en corse⁴⁹ (« Wikiquote ») a été lancée en 2005. Elle compte 6 pages de citations en juin 2014 et moins de 500 contributeurs dont aucun n'a été actif au cours des 30 derniers jours.

Une version de Wikibooks en corse⁵⁰ (« Wikilibri ») a été lancée en 2005, mais fermée à la suite d'une décision prise en 2011. En effet, cette version ne s'était enrichie que de 12 pages en plus de cinq ans. Elle demeure néanmoins toujours consultable.

4.3.9 Flamand occidental

La version de Wikipédia en flamand occidental⁵¹ (« Wikipedia ») a été ouverte en mars 2006, en utilisant le code **vls**.

Elle compte en juin 2014 près de 5 000 articles et plus de 10 000 contributeurs inscrits, dont 19 ont été actifs au cours des 30 derniers jours, soit l'équivalent d'environ 5 contributeurs pour un million de locuteurs.

La page d'accueil de Wikipédia en flamand occidental est consultée en moyenne environ 300 fois par jour. Durant les trois premières semaines du mois de juin 2014, les dix articles les plus consultés, hors pics de consultation ponctuels non représentatifs, étaient « Mots typiques de flamand occidental », « Flamand occidental », « Croatie », « Wikipédia », « Belgique », « France », « États-Unis d'Amérique », « Islam », « Canada » et « Anglais ».



WIKIPEDIA
Den vryen encyclopedie

⁴⁵ <https://co.wikipedia.org>

⁴⁶ <https://co.wiktionary.org>

⁴⁷ https://wikisource.org/wiki/Main_Page/Corsu

⁴⁸ « Dichjarazioni univirsali di i diritti di l'omu » sur

https://wikisource.org/wiki/Dichjarazioni_univirsali_di_i_diritti_di_l'omu

⁴⁹ <https://co.wikiquote.org>

⁵⁰ <https://co.wikibooks.org>

⁵¹ <https://vls.wikipedia.org>

4.3.10 Normand

La Wikipédia en normand⁵² (« Viqùipédie ») a été ouverte en mars 2006, en utilisant le code *nrm*. Il n’y avait pas – et il n’y a toujours pas – de code ISO 639 pour le normand ; le code ISO 639-3 *nrm* désigne le narum ou narom, langue austronésienne parlée en Malaisie.

Le normand regroupant plusieurs dialectes, la Wikipédia en normand accepte au moins les dialectes cotentinais, jersiais et guernesiais. Les articles comportent souvent un bandeau en haut de l’article pour indiquer le dialecte utilisé dans l’article.

Elle compte en juin 2014 plus de 3 500 articles et plus de 6 000 contributeurs inscrits, dont 11 ont été actifs au cours des 30 derniers jours.



VIQÙIPÉDIE
L'encyclopédie libe

La page d’accueil de Wikipédia en normand est consultée en moyenne environ 300 fois par jour. Durant les trois premières semaines du mois de juin 2014, les dix articles les plus consultés, hors pics de consultation ponctuels non représentatifs, étaient « Canada », « Wikipédia », « Normand », « Anglais », « État », « Allemagne », « France », « Royaume-Uni », « Autriche » et « États-Unis d’Amérique ».

Chabot

Ch't' articl'ye est liée doubl'ye : vaie *Chabot (frouque)*. ↩

Ch't' articlle fut c'menchi en **Cotentinais** 🇫🇷

Eun **chabot** (+ **sabot**) est eune caôcheure en boués.

Ole est féchounaée par eun **chaboty** dauns eun bouot de **boués** aisi à creusaer. Les chabots sount déeus manyires. L's sount touot de boués, l's ount ryin qué le pyid dé semèle qu'est en boués, lé desus est eune pyiche de **tchuu** ahoquie ; chute ségounde féchoun 'tait mens chire.

Achteu maisi, i ne fait pus de chabot. Les quenâles ount tous des **caôcheures** pour alaer és **écoles**, les **paisans** ount des **bôtes** dé **câotchou** qui sount pus hâotes et pus durauntes.



Ch't' articlle est eune **chouquette** à abyinaer - Si vos avaez de quèi à ajoutaer, vos pouvaez l'**amendaer**.



Eune pathe dé chabots

*Article « Chabot » sur la Wikipédia en normand – on remarque le bandeau annonçant que l’article en rédigé en dialecte cotentinais.*⁵³

Saint Swithîn

Ch't' articl'ye fut c'menchie en **Jèrriais** ✖

Saint **Swithîn** 'tait un **évêque** dé **Winchester** en **Angliètèrre**.

Duthant sa vie, i' sembl'ye qu'i' n'fit qu'un **mithacl'ye**: eune vieille femme avait laïssi tchaie eune pann'lée d'**oeux**. L'saint fit l'singne d'la **Crouaix** au d'ssus d's oeux crutchis et il' èrd'vîntent entchièrs.

Mais Saint Swithîn est mus connu pour chein tch'arrivit auprès sa mort en l'an **862**. Humbliément, l'**saint** avait c'mandé qu'san corps dév'thait êt' entèrré dans l'**chînm'tchiéthe** d'la **cathédrale** - pas dans eune pliaiche d'hounneu en d'dans - à seule fin qu'sa tombe pûsse êt' arrouôsée par la plyie du ciel.

*Article « Saint Swithîn » sur la Wikipédia en normand – on remarque le bandeau annonçant que l’article est rédigé en jersiais.*⁵⁴

52 <https://nrm.wikipedia.org>

53 Article sous licence CC-BY-SA 3.0 par les contributeurs à l’article « Chabot » sur la Wikipédia en normand. Photo sous licence CC-BY-SA 3.0 par Kuchen sur Wikimedia Commons.

<https://nrm.wikipedia.org/wiki/Chabot>

https://commons.wikimedia.org/wiki/File:Wooden_Shoes-willow-plain_wood.jpg

54 Article sous licence CC-BY-SA 3.0 par les contributeurs à l’article « Saint Swithîn » sur la Wikipédia en normand.

https://nrm.wikipedia.org/wiki/Saint_Swithîn

4.3.11 Picard

La version de Wikipédia en picard⁵⁵ (« Wikipédia ») a été ouverte en 2009, en utilisant le code pcd.

Elle compte en juin 2014 plus de 2 700 articles et plus de 4 600 contributeurs inscrits, dont 16 ont été actifs au cours des 30 derniers jours.

La page d'accueil de Wikipédia en picard est consultée en moyenne environ 350 fois par jour. Durant les trois premières semaines du mois de juin 2014, les dix articles les plus consultés, hors pics de consultation ponctuels non représentatifs, étaient « France », « États-Unis d'Amérique », « Wikipédia », « Picard », « Pologne », « Anglais », « Allemagne », « Russie », « Brésil » et « Afrique ».

Une version du Wiktionnaire en picard⁵⁶ (« Wikamote ») s'élabore sur incubator.wikimedia.org, site dédié aux débuts de futures versions linguistiques de projets Wikimedia. La version compte près de 250 pages en langue picarde en juin 2014.

4.3.12 Francoprovençal

La version de Wikipédia en francoprovençal⁵⁷ (« Vouiquipèdia ») a été ouverte en mars 2006, en utilisant le code frp.

Elle compte en juin 2014 plus de 2 300 articles et plus de 6 400 contributeurs inscrits, dont 15 ont été actifs au cours des 30 derniers jours, soit l'équivalent d'environ 18 contributeurs pour un million de locuteurs.

Comme l'occitan, sur Wikipédia le francoprovençal (ou arpitan) regroupe plusieurs dialectes et patois régionaux, parmi lesquels (indiqués sur le site) : le bressan, le burgondan, le fribourgeois, le lyonnais, le mâconnais, le savoyard, le valaisan, le valdôtain et le vaudois. Pour permettre la cohabitation entre ces différents dialectes, le projet a mis en place une page intitulée « Informations sur la graphie » sur laquelle a été entreprise une démarche visant à « unir [les] écritures pour ne pas trop sombrer dans des petits désaccords orthographiques... ».⁵⁸ Il n'est toutefois pas indiqué sur cette page si une règle claire a finalement été définie sur cette question, bien que l'orthographe de référence B (ORB) y soit mentionnée.⁵⁹

La page d'accueil de Wikipédia en francoprovençal est consultée en moyenne environ 250 fois par jour. Durant les trois premières semaines du mois de juin 2014, les dix articles les plus consultés, hors pics de consultation ponctuels non représentatifs, étaient « Vagin », « Wikipédia », « Italie », « Anglais », « États-Unis d'Amérique », « France », « Arpitan », « Allemagne », « Europe » et « Royaume-Uni ».



WIKIPÉDIA
El libe inciclopédie



WIKIPEDIA
L'enciclopèdia abada

⁵⁵ <https://pcd.wikipedia.org>

⁵⁶ https://incubator.wikimedia.org/wiki/Wt/pcd/Moaite_Pache

⁵⁷ <https://frp.wikipedia.org>

⁵⁸ « Informations sur la graphie », https://frp.wikipedia.org/wiki/Vouiquipèdia:Enformacions_sur_la_grafia

⁵⁹ Le site arpitania.eu indique à propos de l'ORB : « L'ORB est juste une orthographe, qui permet donc de donner une forme unique à des mots qui ont la même étymologie mais des réalisations phonétiques différentes. En aucun cas l'ORB ne sert à uniformiser la syntaxe des parlers des différentes zones de l'Arpitanie, ni même la prononciation. » <http://www.arpitania.eu/index.php/langue-arpitan-francoprovençal/2-arpitania/58-arpitan-francoprovençal-orthographe-de-reference-b>

5 . Conclusion

En conclusion, les langues de France hexagonales possèdent en 2014 une place qui n'est pas négligeable sur les projets Wikimedia, avec un total de plus d'un million de pages de contenu existant en juin 2014.

Concernant plus particulièrement le projet Wikipédia, ajoutées à la langue française, l'ensemble des langues de France ne représentent que la moitié des articles existant en français (845 000 articles de Wikipédia en langues de France – sans le français – contre 1 515 000 articles en français). Pour autant, certaines s'avèrent plus actives que la Wikipédia francophone eu égard du nombre de contributeurs actifs pour un million de locuteurs, ce qui est le cas du breton, du basque et du catalan.

Par ailleurs, à l'exception du breton, toutes les langues présentes sont des langues situées près de frontières et sont donc également parlées dans d'autres pays, le plus souvent frontaliers. Cela signifie qu'en dehors de ces zones frontalières ainsi que de la zone couverte par l'occitan au sud du pays, une autre grande zone, s'étendant de l'Atlantique à la Lorraine et comprenant majoritairement des langues d'oïl, ne compte aucune langue ni dialecte sur les projets Wikimedia. En effet, en l'absence d'une reconnaissance par un code ISO 639, ces parlers sont exclus des projets Wikimedia et plus généralement du monde numérique. La DGLFLF pourrait ici apporter une aide précieuse pour améliorer la reconnaissance à l'international de ces langues.

Les niveaux de consultation des projets Wikipédia en langues de France sont également significatifs, avec pour le projet le moins consulté, Wikipédia en francoprovençal, environ 250 consultations journalières de la seule page d'accueil. À l'inverse, le projet Wikipédia le plus consulté, en langue catalane, compte en moyenne 10 000 consultations journalières de sa page d'accueil. Les articles les plus consultés indiquent par ailleurs que la plupart des visiteurs s'intéressent aux articles traitant de Wikipédia, de leur langue régionale, ou de Wikipédia en leur langue régionale. Apparaissent aussi souvent la France, les États-Unis, l'Allemagne et le Royaume-Uni, ainsi que la langue anglaise. Enfin, l'actualité marque aussi une partie des consultations avec en tête au mois de juin 2014 les articles sur la Seconde Guerre mondiale, la Coupe du monde de football au Brésil ou la nuit de la Saint-Jean, tout en sachant que la présence des États-Unis, du Royaume-Uni et dans une moindre mesure de la Russie et du Canada peuvent aussi être expliqués par les commémorations du 70^e anniversaire du Débarquement de Normandie.

Enfin, à l'exception d'une poignée de versions restant peu voire pas du tout modifiées faute de contributeurs suffisants, la majorité des projets Wikimedia en langues de France hexagonales demeurent malgré tout relativement actifs.

Ce rapport est placé sous licence CC-BY-SA 3.0 par ses contributeurs Florian Pépelin, Nicolas Vignerot, Sébastien Beyou, Pierre-Antoine Le Page, Caroline Becker, Guillaume Goursat. Les logos des projets Wikimedia sont des marques déposées de la Wikimedia Foundation.