

20 Jahre kleine Sprachversionen – ein Zwischenstand

Holder

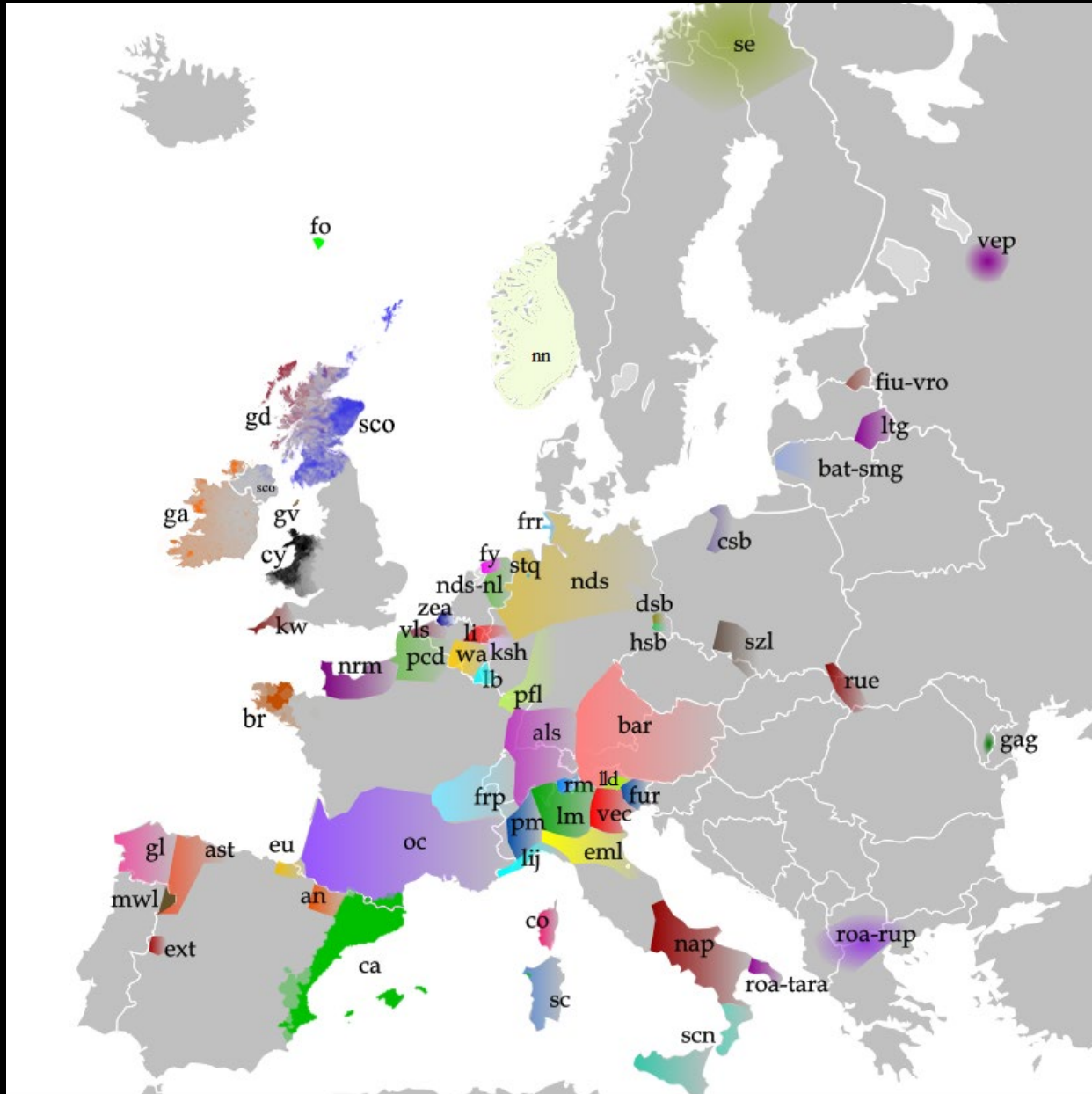
WikiCon 2023, Linz, 30.09.2023



Klein, aber fein ...

... oder nicht doch eher klein und hässlich?

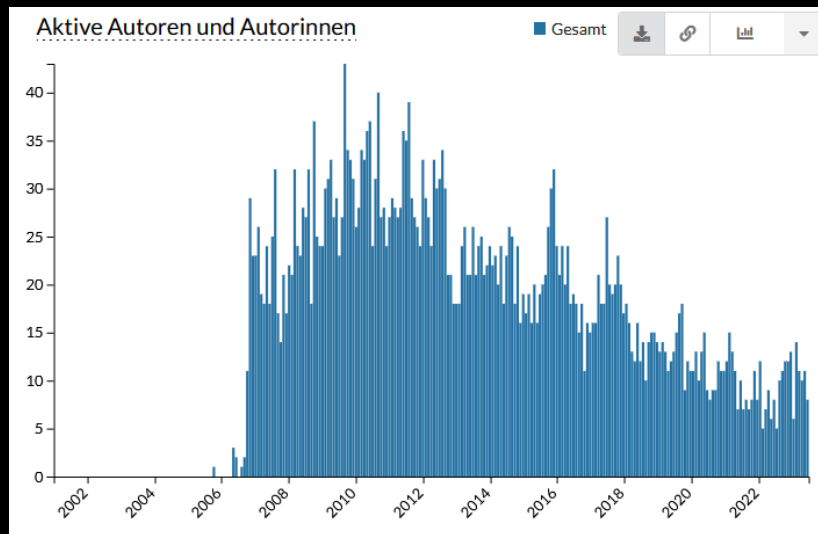
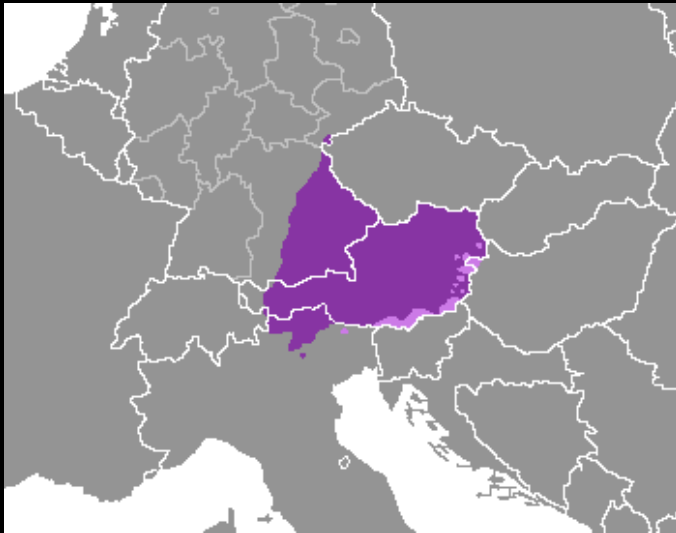
Welche Sprachen?



Jimmy Wales, 2005

“Wikipedia is first and foremost an effort to create and distribute a free encyclopedia of the highest possible quality to every single person on the planet in their own language.”

Steckbriefe: bar:wp



26 969 Artikel
(seit 09/2006)

2 Admins

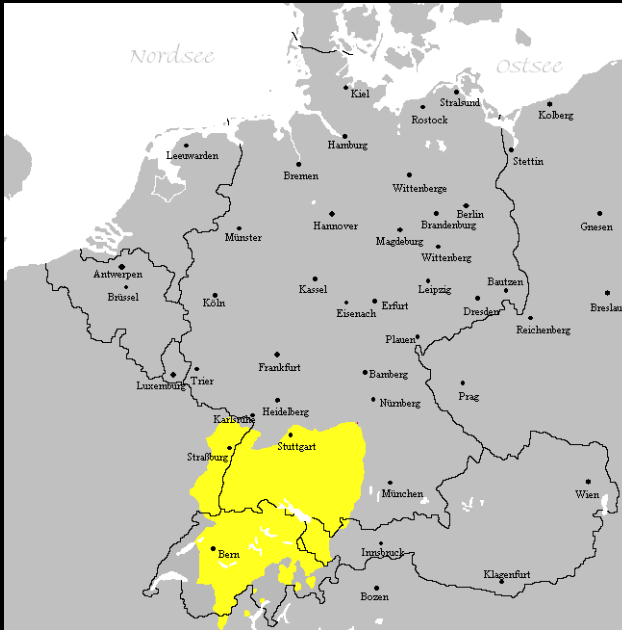
214 Wörter pro
Artikel

Jahrelanger
Streit um
Varietäten und
Rechtschreibung

Community im
Streit zerbrochen

Mittlerweile fast
inaktiv

Steckbriefe: als:wp



29 906 Artikel
(seit 11/2003)

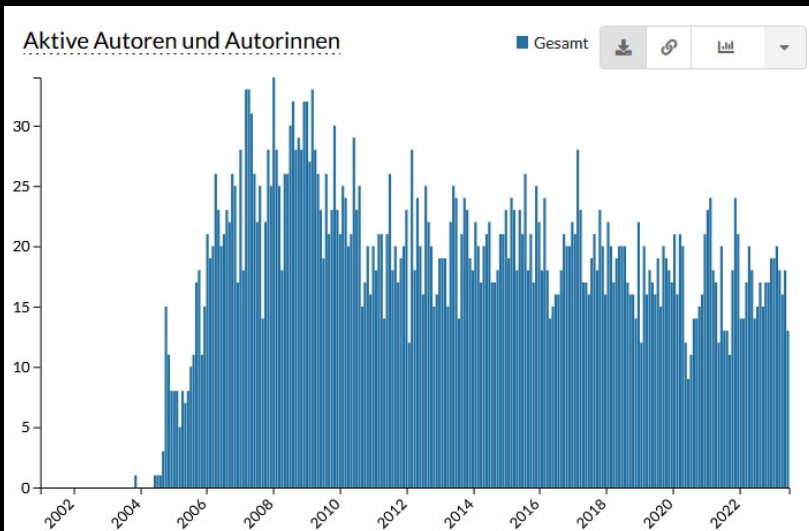
1 036 398 Edits
insgesamt

436 Wörter pro Artikel

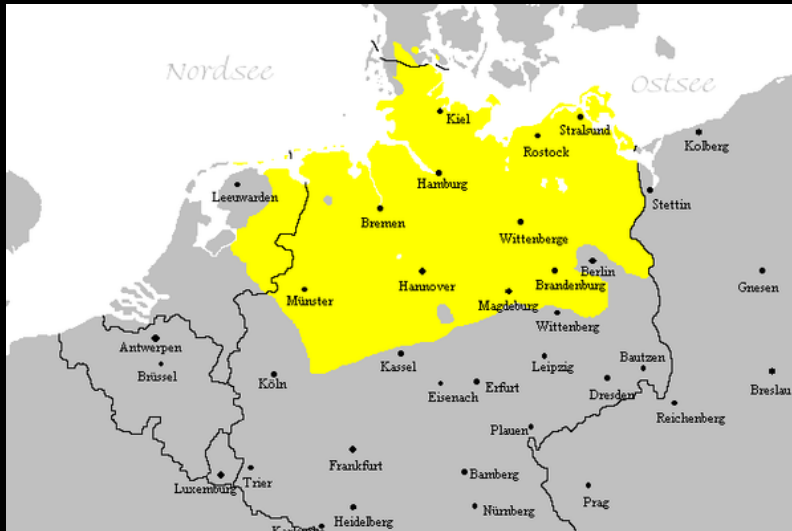
Startseite in fünf
verschiedenen
Varietäten:
Schwyzerdütsch,
Badisch, Schwäbisch,
Elsassisch,
Vorarlbergisch

Alle Dialekte und
Orthographiesysteme
werden toleriert.

Schwesterprojekte als
eigene Namensräume
in Wikipedia:
Wikisource, Wikinews,
Wiktionary, Wikibooks,
Wikiquote



Steckbriefe: nds:wp



84 251 Artikel
(seit 04/2003)

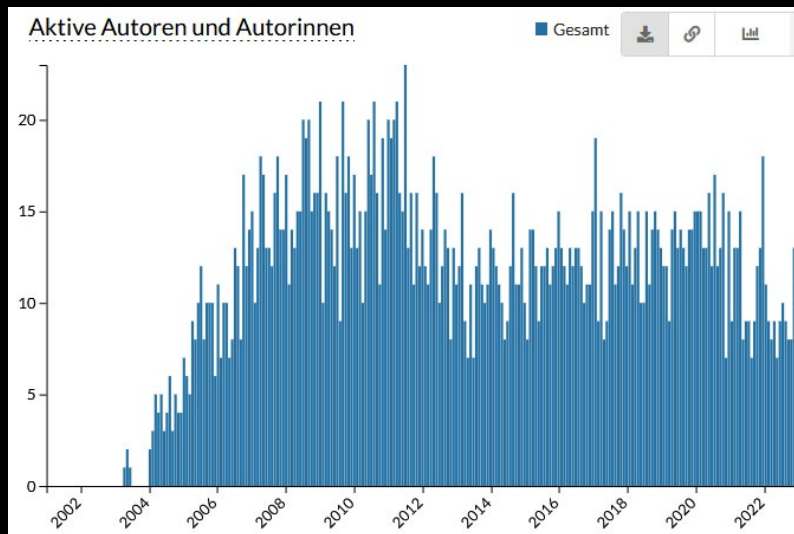
davon über 50.000
BKL (> 60 % der
Seiten im ANR)

186 Wörter pro
Artikel

Nur Orthografie
nach Sass
„Plattdeutsches
Wörterbuch“ wird
toleriert.

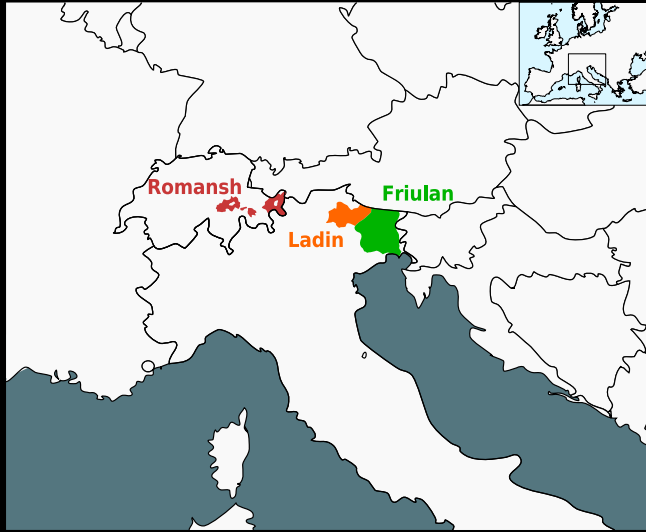
Deshalb
Abspaltung der
Wikipedia in
Nedersaksisch
(nds-nl) 2006

Im Incubator:
Wikipedia in
Plautdietsch



nds:wiktionary
umfasst alle drei
Varietäten

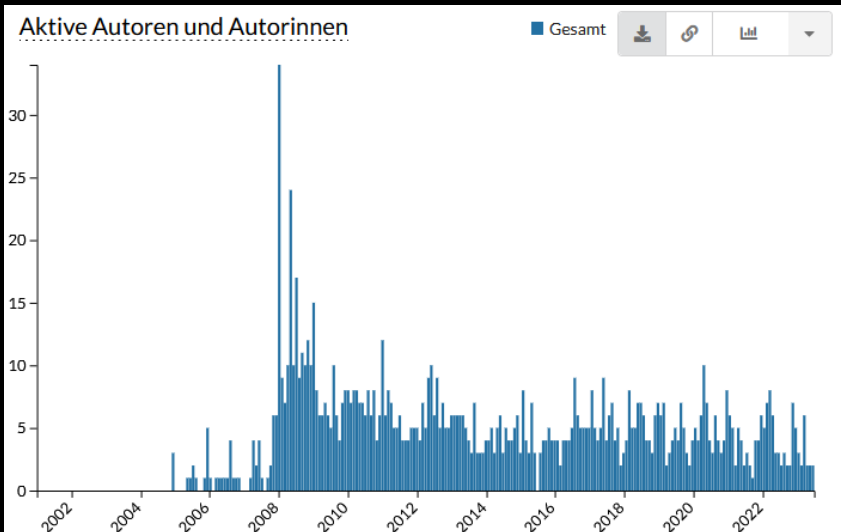
Steckbriefe: rm:wp



Rumantsch Grischun
und traditionelle
Idiome

3 764 Artikel
(seit 12/2003)

164 341 Edits
insgesamt



Laut offizieller
Statistik praktisch
inaktiv

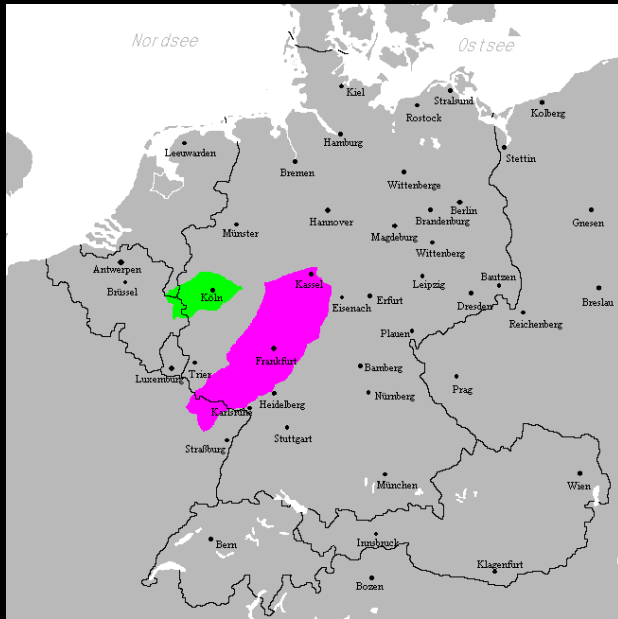
und seit Jahren keine
Diskussionen

Aber:
Regelmäßige Übersetzungen exzellenter und
lesenswerter Artikel aus der de:wp

309 Artikel sind größer als 20
kB,
23 sind größer als 100 kB

756 Wörter pro
Artikel

Steckbriefe: ksh:wp und pfl:wp



ksh:
2 937 Artikel
1 607 612 Edits
(seit 11/2005)

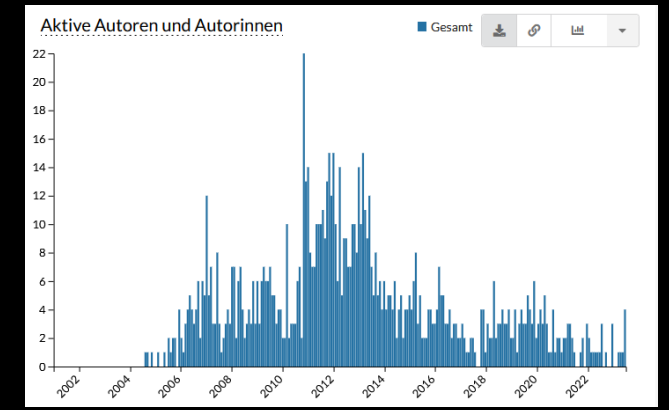
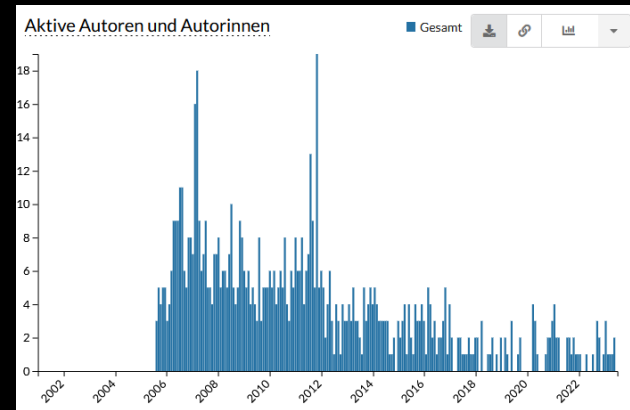
pfl:
2 745 Artikel
90 631 Edits
(seit 11/2010)

Hohe Toleranz
gegenüber
Varietäten und
Schreibweisen

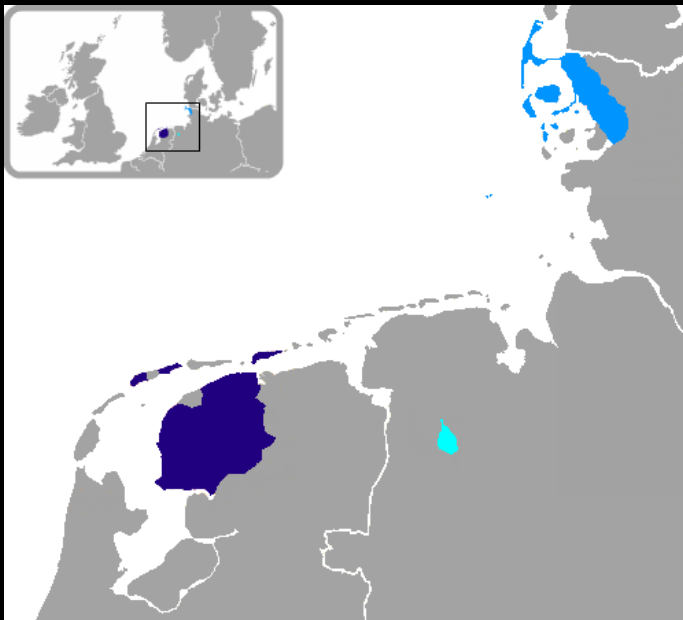
ksh:
Community im
Streit zerbrochen

pfl
Nach Import aus
Incubator recht
aktiv, aber bald
wieder
eingeschlafen

Heute beide
Projekte praktisch
inaktiv



Steckbriefe: fy:wp, frr:wp und stq:wp



fy:
52 115 Artikel
(seit 02/2002)

frr:
17 821 Artikel
(seit 08/2010)

stq:
4 095 Artikel
(seit 01/2008)

fy:
12-15 regelmäßig
Aktive

frr:
2-3 regelmäßig
Aktive

stq:
mittlerweile praktisch
inaktiv, lange Jahre 4
Aktive

drei friesische Sprachen:
Westfriesisch, Nordfriesisch, Saterfriesisch

Schwesterprojekte

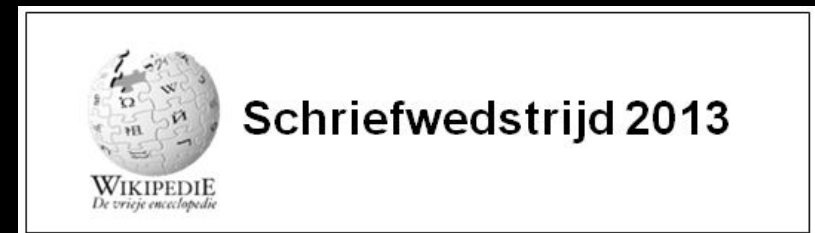
Schwesterprojekte als eigene Namensräume in Wikipedia

- als: Wikisource, Wikinews, Wiktionary, Wikibooks , Wikiquote
- bar: Wikisource, Wikinews, Wiktionary, Wikibooks , Wikiquote
- pfl: Wikisource, Wikinews, Wiktionary, Wikibooks , Wikiquote
- frr: Wikisource
- sw: Wikisource



Schreibwettbewerbe

- als: 10 SW seit 2011
- pfl: zwei SW 2014 und 2015
- nds-ndl: ein SW 2009/2010
- zea: ein SW 2013



- z. T. mit Preisen von externen Organisationen (als, nds-nl)
- Bisher keine Schreibwettbewerbe: z. B. nds, bar, ffr, stq, hsb, dsb, rm ...

Artikelauszeichnungen

- als: 195 ausgezeichnete Artikel bei 256 Kandidaturen seit 2006
 - bar: 172 ausgezeichnete Artikel bei 200 Kandidaturen seit 2007
 - ext: 4 ausgezeichnete Artikel bei 6 Kandidaturen seit 2009
 - zea: 7 ausgezeichnete Artikel bei 10 Kandidaturen seit 2014
-
- Manche Projekte (z. B. Afrikaans) wählen so die Artikel für die Startseite aus > regelmäßige Kandidaturen und Auszeichnungen





Treffen

- als: regelmäßige Treffen seit 2006
- lb: regelmäßige Treffen seit 2006 (?)
- bar: ein Treffen 2007
- ksh: ein Treffen 2007
- pfl: ein Treffen 2010
- nds: Treffen diskutiert, sind aber nie zustande gekommen ...
- Teilnahme an WikiCon (de), WikiConvention (fr), Wikimania, CEE usw.

Celtic Knot Conference

- Celtic Knot Conference 2017 (Edinburgh, Scotland)
- Celtic Knot Conference 2018 (Aberystwyth, Wales)
- Celtic Knot Conference 2019 (Penryn, Cornwall)
- Celtic Knot Conference 2020 (online, originally planned for Limerick, Ireland)
- Arctic Knot Conference 2021 (online, originally planned for Tromsø, Norway)
- Celtic Knot Conference 2022 (online + local gatherings)
- 2023: no conference this year.



Pilotprojekt Atikamekw (atj)

- Atikamekw (unterstützt durch Wikimedia Canada, gefördert durch WMF):
- Wikimedians of North American Indigenous Languages User Group
- Pilotprojekt in einer Sprache der First Nations
- Erstes Projekt 2013/2014, zweites Projekt 2016-2021
 - Workshops
 - Aktuelle Dokumentation von traditionellen Aktivitäten
 - Bilder und Dokumente aus Archiven auf Commons hochgeladen
 - Unterstützung beim Aufbau einer Wikipedia in Atikamekw



Interessantes

- Diné bizaad (nv): Navajo-Community wünscht sich keine Artikel in der WP zur traditionellen Religion
- Schlesisch (sli, noch im Incubator): WP beschränkt sich auf Themen rund um Schlesien
- Tongaisch (to): Artikel zu Tonga zusätzlich mit englischer Übersetzung
- Walisisch (cy): Bei Personenartikeln sind 50 % Artikel über Frauen.
- Nordfriesisch (frr): Artikel sind z. T. in mehreren Sprach-Varietäten vorhanden.
- Die beiden niederdeutschen Wikipedien (nds und nds-nl) hatten viele Jahre auf ihrer Startseite neben ihrem Artikel des Monats auch jeweils den der andern Sprachversion.

Herausforderungen und Probleme

- Kaum Vandalismus, man freut sich geradezu über Vandalismus in der jeweiligen Sprache!
- Fast keine Werbeeinträge, fast keine Einflussnahme von außen.
- Neue Artikel schreiben ist attraktiver als vorhandene pflegen > Artikel veralten mit der Zeit.
- Regelmäßige Updates der Software machen Pflege des Projekts schwierig.
- Vorlagen veralten mit der Zeit.
- Stewards erkennen Adminwahlen z. T. nicht an oder vergeben nur befristete Rechte.

Ernüchterndes

- Schottisch (sco): Nicht-Muttersprachler schreibt über 30.000 Artikel in Pseudo-Schottisch > mittlerweile davon 2/3 wieder gelöscht.
- Nördliches Lurisch (lrc): Benutzer schreibt Artikel in Sprache, die für Muttersprachler unverständlich ist > Projekt geschlossen, alle Artikel gelöscht.
- Lombardisch (lom): 2 Benutzer schreiben über 100.000 Artikel in erfundener Standardsprache > alle diese Artikel gelöscht.
- Ripuarisch (ksh): 10.000 Artikel zu Postleitzahlen und zu Telefonvorwahlen > alle diese Artikel gelöscht.

Bot-generierte Artikel

- Walisisch (cy): Per Bot aktualisierte Listen aus WikiData (in 181.743 von 279.352 Artikeln)
- Plattdeutsch (nds): 50.000 BKL per Skript ohne Bot-Account (> 60 % der Seiten im ANR)
- Asturianisch (ast): Automatisierte Übersetzungen aus dem Spanischen > Sprachliche Qualität?
- Ladinisch (lld): 178.000 von 180.541 sind Kurzaufartikel (1-2 Sätze) per Skript ohne Bot-Account
- Ägyptisch-Arabisch (arz): 1.579.881 von 1.619.298 sind Kurzaufartikel per Skript ohne Bot-Account
- Aromunisch (roa-rup): 50.000 bot-generierte Artikel, mittlerweile wieder gelöscht
- Okzitanisch (oc): 50 000 bot-generierte Artikel 2011-2013, davon 10.000 wieder gelöscht.
- Tatarisch (tt): 450.000 von 500.000 Artikeln (90 %)
- Cebuano (ceb): 6.073.157 von 6.122.996 Artikeln (99 %)

Fazit

- Mit den großen Sprachversionen und deren Inhalten und Communitys können die kleine Sprachgemeinschaften nicht verglichen werden.
- In vielen Fällen bleibt es ein Experiment.
- Kaum eine kleine Sprachversion konzentriert sich auf die eigene Region und Kultur.
- Leider oft eher Quantität statt Qualität.
- Bot-generierte Artikel, automatische Artikel aus Wikidata, Abstract Wikipedia o. ä. helfen den kleinen Sprachen nicht weiter.
- Wenn sich engagierte Leute finden, kann aber durchaus ein interessantes Projekt mit sinnvollen Inhalten entstehen!

Merci viilmool!



**Kein englischer
Rasen, sondern eine
bunte Blumenwiese!**

Bildquellen

- https://commons.wikimedia.org/wiki/File:Charte_Wikipedia_in_Regionalsproche.png (Holder, CC BY-SA 3.0)
- https://commons.wikimedia.org/wiki/File:European_Wikipedias_article_count_2019_map.svg (Dexxor, CC BY-SA 4.0)
- https://commons.wikimedia.org/wiki/File:Dialecto_austro-bávaro.PNG (Fobos92, CC BY-SA 4.0)
- https://commons.wikimedia.org/wiki/File:Schwäbisch-Alemannisches_Mundartgebiet.PNG (Postmann Michael, PD)
- https://commons.wikimedia.org/wiki/File:Low_Saxon_language_area.png (Postmann Michael, PD)
- https://commons.wikimedia.org/wiki/File:Rhaeto-Romance_languages.svg (ArnoldPlaton, CC BY-SA 3.0)
- https://commons.wikimedia.org/wiki/File:Zone_des_francique_rhénan_et_ripuaire.png (Odejea, GNU FDL 1.2)
- https://commons.wikimedia.org/wiki/File:Frisian_languages_in_Europe.svg (ArnoldPlaton, CC BY-SA 3.0)
- https://commons.wikimedia.org/wiki/File:Wiki-Treffen_Mee_2015.jpg (MMFE, CC BY-SA 3.0)
- https://commons.wikimedia.org/wiki/File:Igbo_Wiki_Meetup_group_picture.jpg (Uzoma Ozurumba, CC BY-SA 4.0)
- https://commons.wikimedia.org/wiki/File:Celtic_Knot_Conference_2018_at_NLW_23.jpg (Jason.nlw, CC0 1.0)
- https://commons.wikimedia.org/wiki/File:ManawanOct2016_14.jpg (Pierre Coulombe, CC BY-SA 4.0)
- <https://commons.wikimedia.org/wiki/File:Gasterental.jpg> (Daniel Baumgartner, CC-BY-SA-2.5)