

Holder

Botpedia

WikiCon 2014, Köln 3.-5.10.2014

Botpedia

Skriptgenerierte Artikel

Vergangenheit, Gegenwart und Zukunft der
vielsprachigen Wikipedia?

Eine Reise durch gewünschte Höhen und
reale Tiefen der Artikelprogrammierung

Typologie

- Bot-generierte Artikel
- Skript-generierte Artikel
- Skript-generierte Artikel ohne Bot
- Bot-generierte Artikel ohne Skript
- Massenstubs ohne Skript und ohne Bot

Methoden I

Automatisierte Übersetzungen

- Beispiel: Asturianische Wikipedia (Tradubot)
- http://ast.wikipedia.org/wiki/Recursu_natural
- http://es.wikipedia.org/wiki/Recurso_natural

Methoden II

Datenbanken als Quelle

- Veraltete Informationen
- Falsche Informationen

Data-Mining

Unter **Data-Mining** ['deɪtə 'maɪnɪŋ] (englisch für *Daten-Bergbau*, sinngemäß „in einem Datenberg nach wertvollem Wissen suchen“) versteht man die systematische Anwendung statistischer Methoden auf einen Datenbestand mit dem Ziel, *neue* Muster zu erkennen. Hierbei geht es auch um die Verarbeitung sehr großer Datenbestände (die nicht mehr manuell verarbeitet werden könnten) [...]

Historischer Überblick I

- Rambot auf enwiki, Gemeinden USA, 2002-2004
- [http://en.wikipedia.org/w/index.php?title=Emelle, Alabama&oldid=2947962](http://en.wikipedia.org/w/index.php?title=Emelle,_Alabama&oldid=2947962)
- Projekt auf meta: „Botopedia“, 2005/06
- Projekt auf meta: „Mass content adding“, 2006

Historischer Überblick II

Lombardische Wikipedia

- rund 100.000 Bot-Artikel wieder gelöscht, 2007

Volapük-Wikipedia

- Diskussion über Schließung des Projekts, 2007/08
- https://vo.wikipedia.org/wiki/La_Chapelle-Montlinard

Schema

- [a] ist eine Gemeinde in der Provinz [b] im Land [c]. Die Einwohnerzahl lag im Jahr [d] bei [e].

Historischer Überblick III

- Französische Gemeinden (36.000), z. B. cawiki, euwiki
- https://ca.wikipedia.org/wiki/La_Chapelle-Montlinard
- ca, ceb, es, eu, it, la, ms, oc, pl, pms, pt, ro, ru, sh, sk, sr, sv, uk, uz, vi, war
- Gemeinden in Italien, Spanien, Indien, usw.

Tier- und Pflanzenarten

- Nlwiki: Tierarten, 2011-1013
- Lsjbot (sv, ceb, war): Tierarten, 2012/2013
- https://nl.wikipedia.org/wiki/Chaetolopha_turbinata

Diskussion nl:wp

- Abstimmungen nach Diskussionen
 - Bot-Artikel als Ausgangsbasis für Erweiterung
 - Arbeitserleichterung durch Übernahme von lästigen Arbeiten
 - Themengebiete voll durchstrukturiert mit blauen Links, Koordinaten, z. B. alle Verwaltungseinheiten
- Bedingungen:
 - 3 Fakten
 - Artikel müssen verlinkt sein.
 - Botbetreiber sind verantwortlich für Integration in Struktur, Aktualisierung, Fehlerkorrektur

Seen

Naskobot (sv:wp)

- [http://sv.wikipedia.org/wiki/Örtjärn, Västergötland](http://sv.wikipedia.org/wiki/Örtjärn,_Västergötland)

Botartikel à la Wikidata

- Okzitanische Wikipedia
- [https://oc.wikipedia.org/wiki/Adèle de França](https://oc.wikipedia.org/wiki/Adèle_de_França)
- Projektinterne Diskussion
 - Bot-Betreiber nicht bereit zu Diskussion
 - Arrogantes Auftreten von außen
 - Versuch einer langfristigen Lösung

<http://tools.wmflabs.org/reasonator/?q=Q1339>

Pfäffikon SZ

„Ortschaft der Gemeinde Freienbach im Kanton Schwyz, Schweiz“
(Meyers Universallexikon in 12 Bänden, 2012)

- [http://uz.wikipedia.org/wiki/Pfäffikon \(Schwyz\)](http://uz.wikipedia.org/wiki/Pfäffikon_(Schwyz))

Zukunft

Lsjbot

- Artikel zu Tier- und Pflanzenarten Wikipedia, Arabische Wikipedia

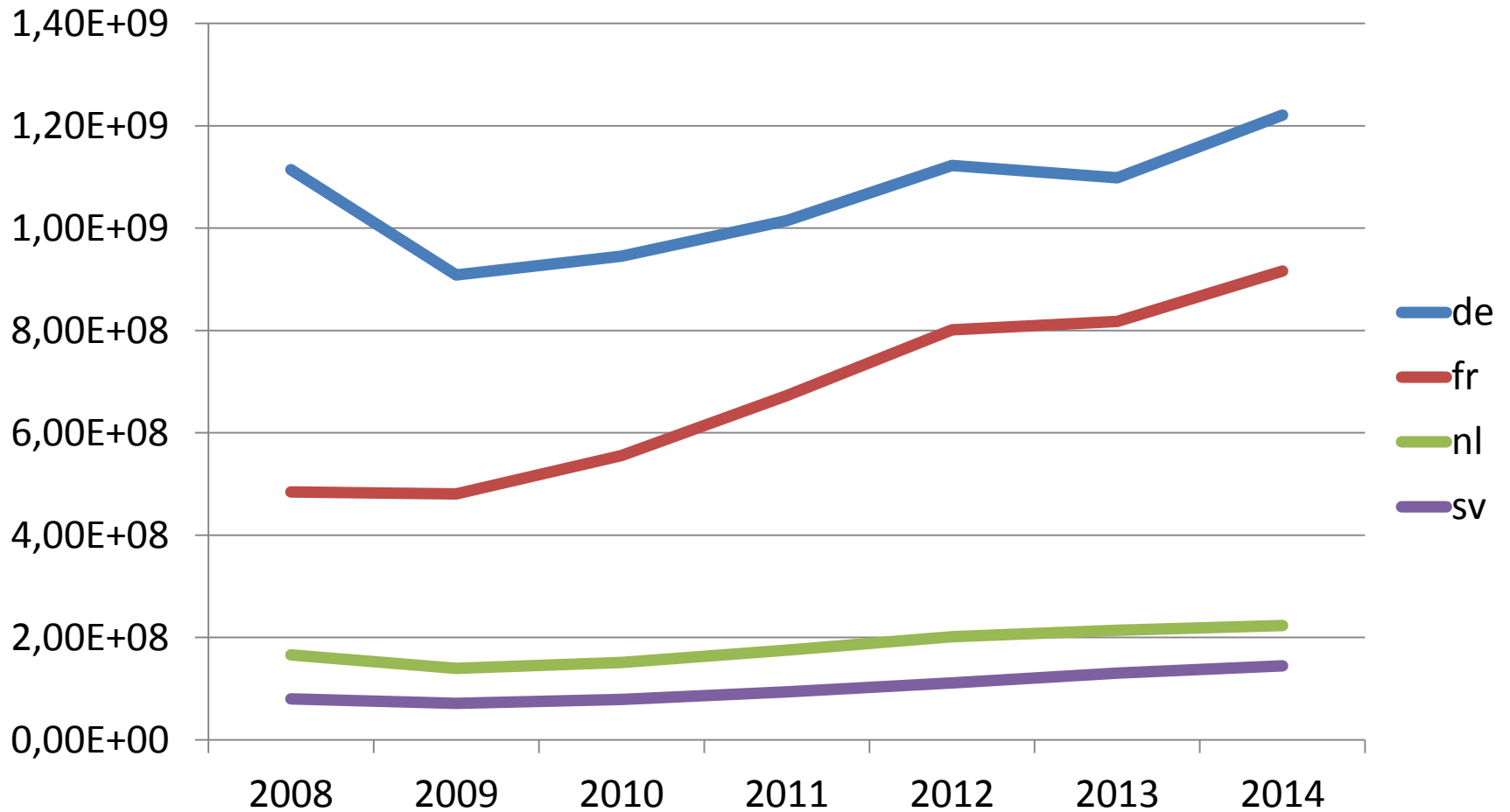
Pläne in verschiedenen Sprachversionen:

- Geographische Objekte
- Bücher
- Denkmäler
- Astronomische Objekte

Selbstverständnis der Wikipedia

- Freies Wissen
- Mehr als andere Internetseiten
- Lexikonartikel: Kurzartikel oder Monografien?

Schrecken Stubs Leser ab?



Schrecken Stubs Autoren ab?

- <http://stats.wikimedia.org/EN/ChartsWikipediaNL.htm>
- <http://stats.wikimedia.org/EN/ChartsWikipediaSV.htm>
- <http://stats.wikimedia.org/EN/ChartsWikipediaDE.htm>