Semiautomated editing in Wikipedia: The AutoWikiBrowser case
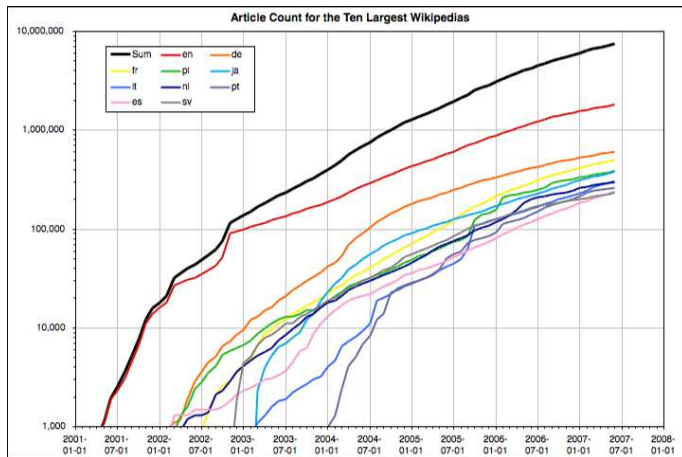
Marios Magioladitis
Sysop in en.wiki,
editor in ar.wiki and de.wiki,
autoreviewer and rollbacker in pt.wiki
autoeditor in ru.wiki,
autopatrolled in fr.wiki
autoconfirmed in es.wiki, tr.wiki and zh.wiki

Carl von Ossietzky Universität Oldenburg, Institut für Mathematik

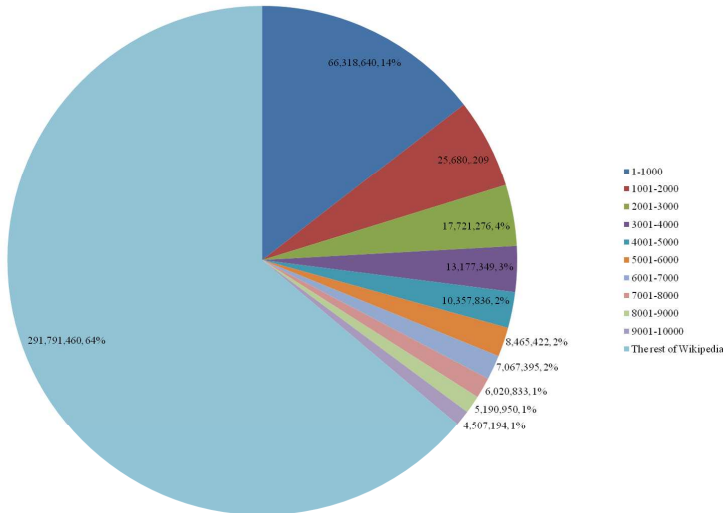15. September 2011
Doktorandenkolloquium

Wikipedia

- Wikipedia founded 15th of January 2001
- 282 languages / 365 million visitors

Article Count for the Ten Largest Wikipedias

- ALL: 19 M / EN: 3,7 M / DE: 1,2 M

The Top 10,000 Compared to the Rest of the Community

66,318,640, 14%

25,680, 209

17,721,276, 4%

13,177,349, 3%

10,357,836, 2%

8,465,422, 2%

7,067,395, 2%

6,020,833, 1%

5,190,950, 1%

4,507,194, 1%

291,791,460, 64%

- 1-1000
- 1001-2000
- 2001-3000
- 3001-4000
- 4001-5000
- 5001-6000
- 6001-7000
- 7001-8000
- 8001-9000
- 9001-10000
- The rest of Wikipedia

Active Wikipedians

The free encyclopedia that anyone can edit

- We still need some rules both on context and layout
- Wikipedia:Wie gute Artikel aussehen
- Alt+TAB

The free encyclopedia that anyone can edit

- We still need some rules both on context and layout
- Wikipedia:Wie gute Artikel aussehen
- Alt+TAB
- Idea: Use (semi-)automated tools to uniformise the style

AutoWikiBrowser (AWB)

Alt+TAB

AutoWikiBrowser in Wikipedia

- Page edits since Wikipedia was set up: 485,598,110 (Special:Statistics)
- Page edits since Wikipedia was set up with AWB: 120,782,035 (toolserver.org/AWB)
- Edits with AWB / Total edits: 24,87%

AutoWikiBrowser: Features

- WP:GENFIXES

Alt+TAB

Regular expressions

- Wikipedia:Regular expressions

Alt+TAB

Code analysis

- > 2 million lines of C#
- 633K comments
- 1.6M is C#
- Many plugins
- Code available at SourceForge.net

Developers

- Bluemoose
- Ligulem
- Magioladitis
- Rjwilmsi
- MaxSem
- Reedy

AutoWikiBrowser: Pros, cons and future plans

Pros and cons

- It works for all wikiprojects and wikias
- It's open source
- It works only for Windows

Future plans

- Mono and wine (Alt+TAB)
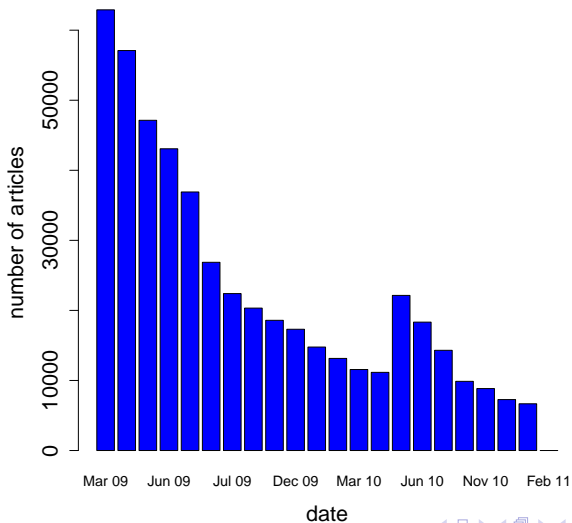- PyAutoWikiBrowser
- Webapp?
- Firefox plugin?

A large example of use: Tagging BLPs

- Wikipedia:Artikel über lebende Personen (Alt+TAB)

- 900k+ biographies in en.wiki
- 60k+ biographies with unknown status in 2009
- Tenths of new entries keep coming daily

- Use prexisting categories
- Always look at the bright side of death
- Create categories by using the text
- Eliminate Dab pages, redirects, etc.
- Extra problems occured by bands etc.
- WP:DUCK

A large example of use: Tagging BLPs

Vielen Dank!

Ευχαριστώ!