



# EDIT & REVERT TRENDS

Erik Zachte

Data Analyst , Wikimedia Foundation

Wikimania

July 2010 - Gdansk





Edits

— Cumulative  
total 556.6 M (100%)

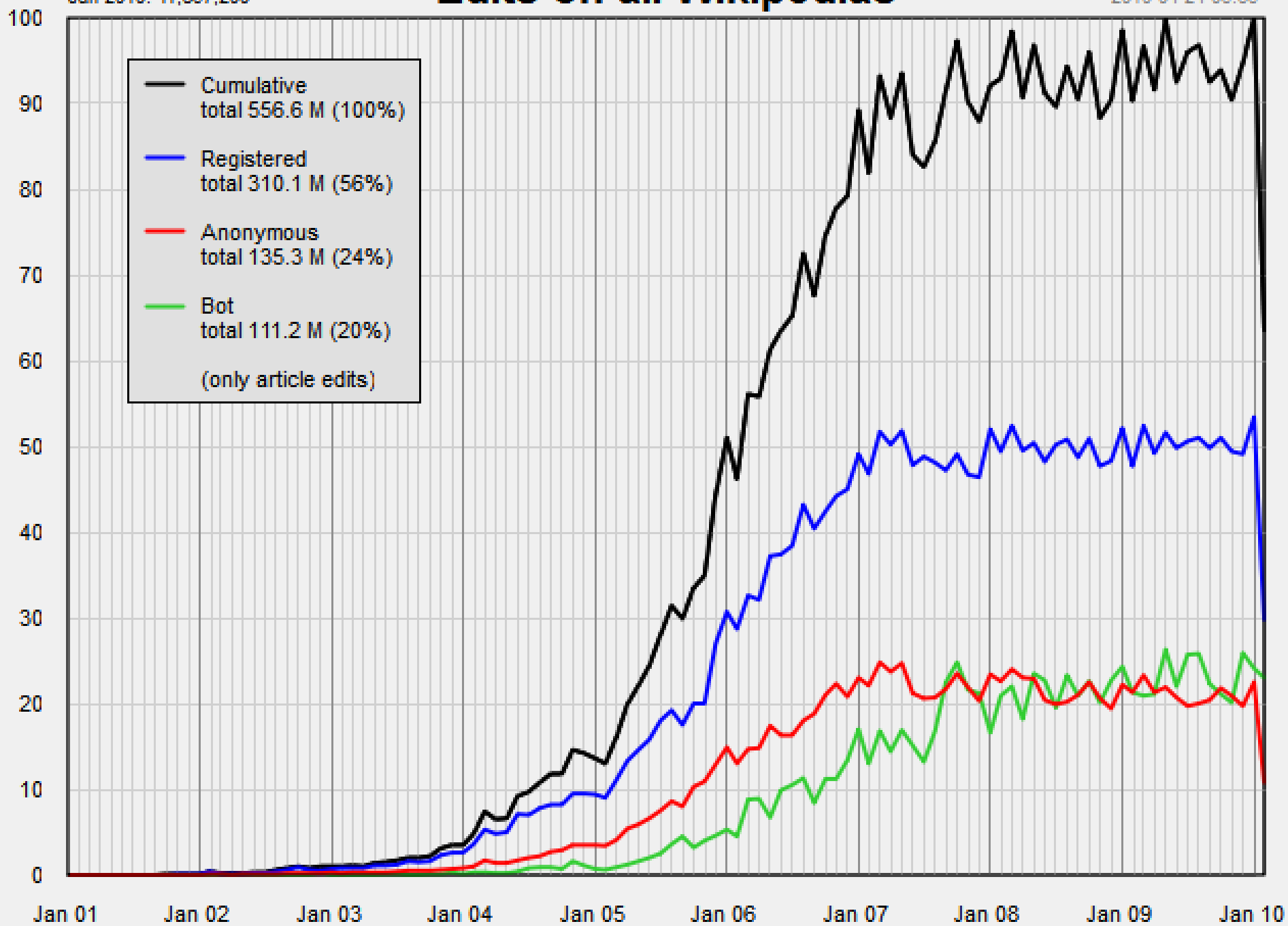
— Registered  
total 310.1 M (56%)

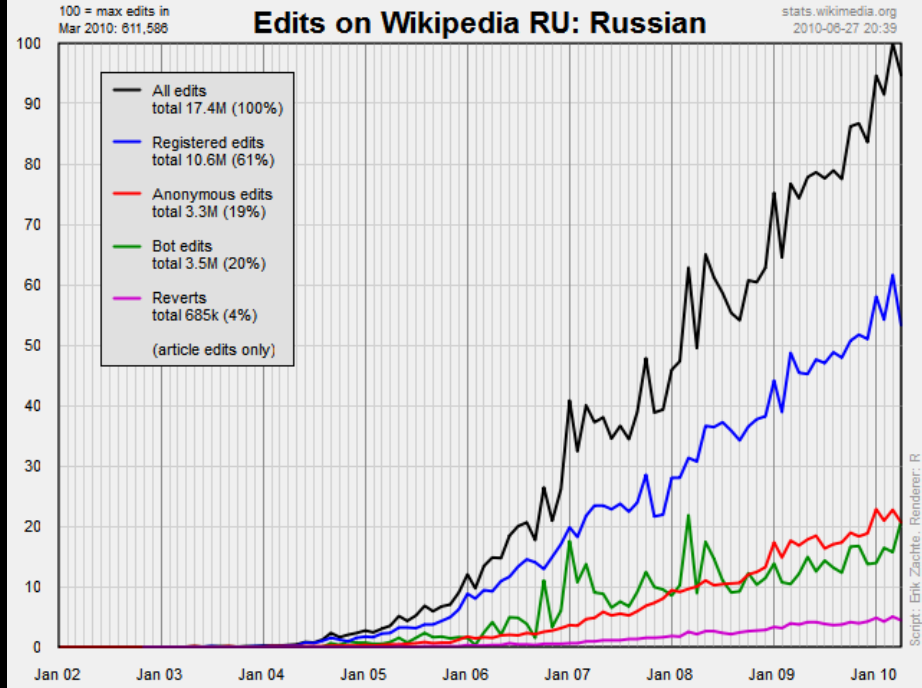
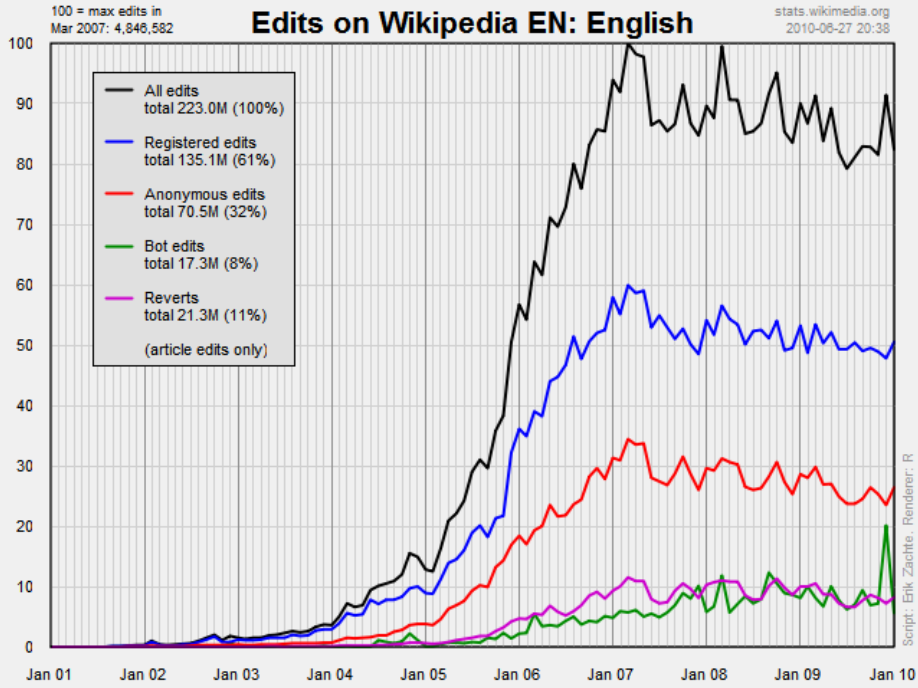
— Anonymous  
total 135.3 M (24%)

— Bot  
total 111.2 M (20%)

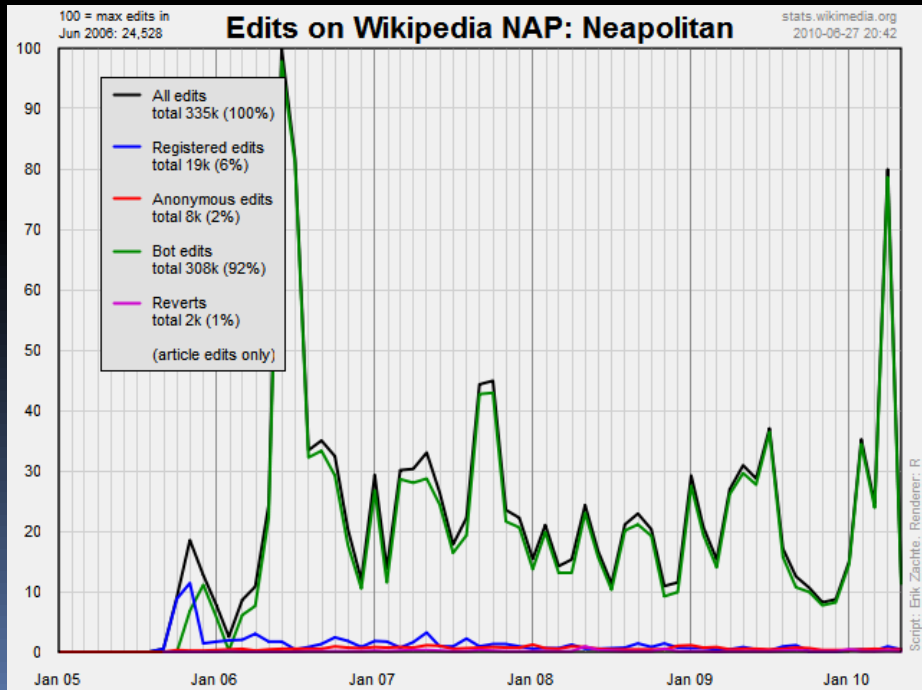
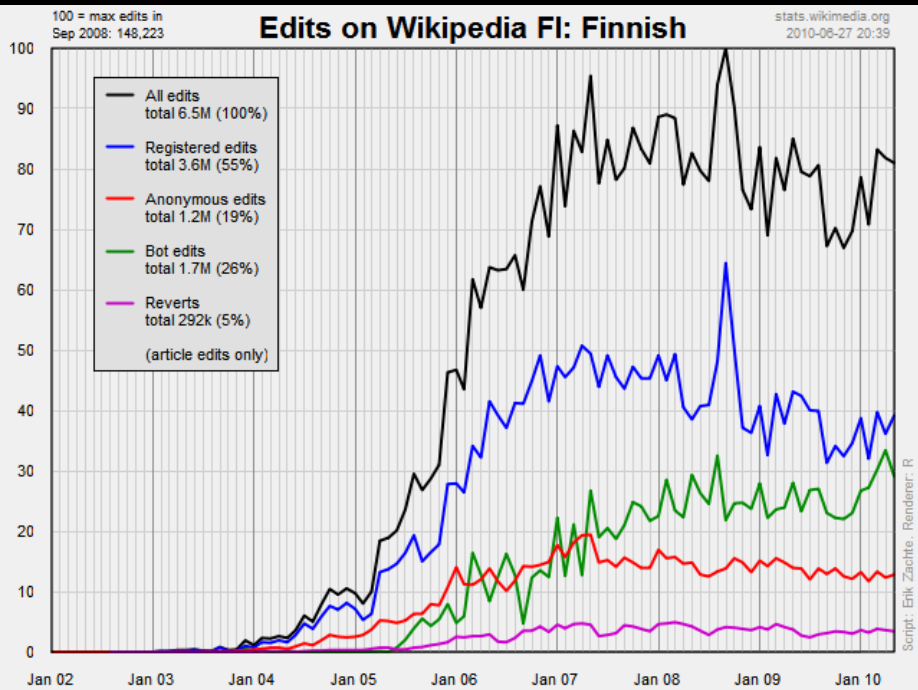
(only article edits)

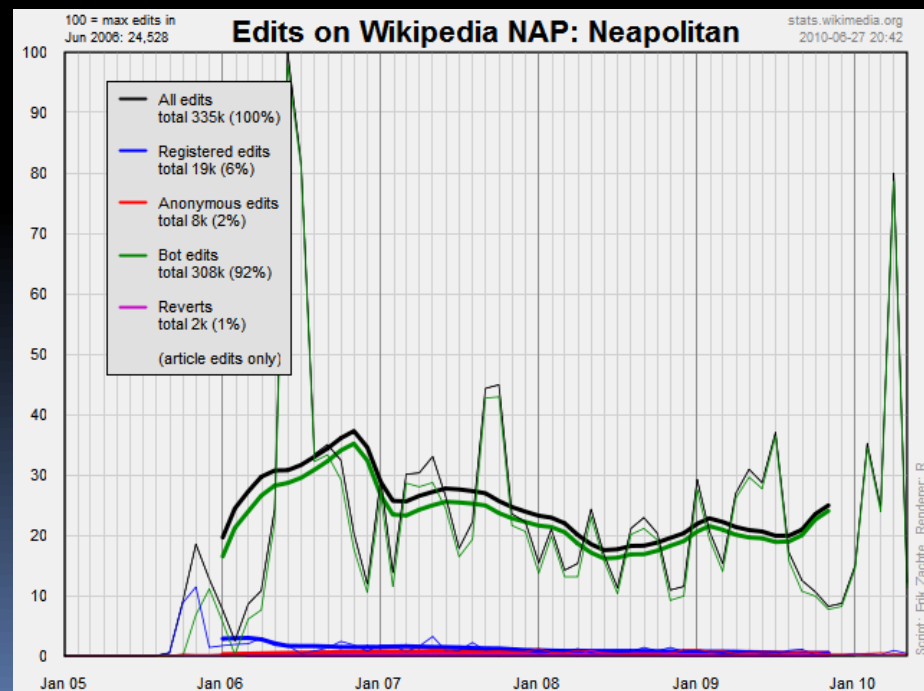
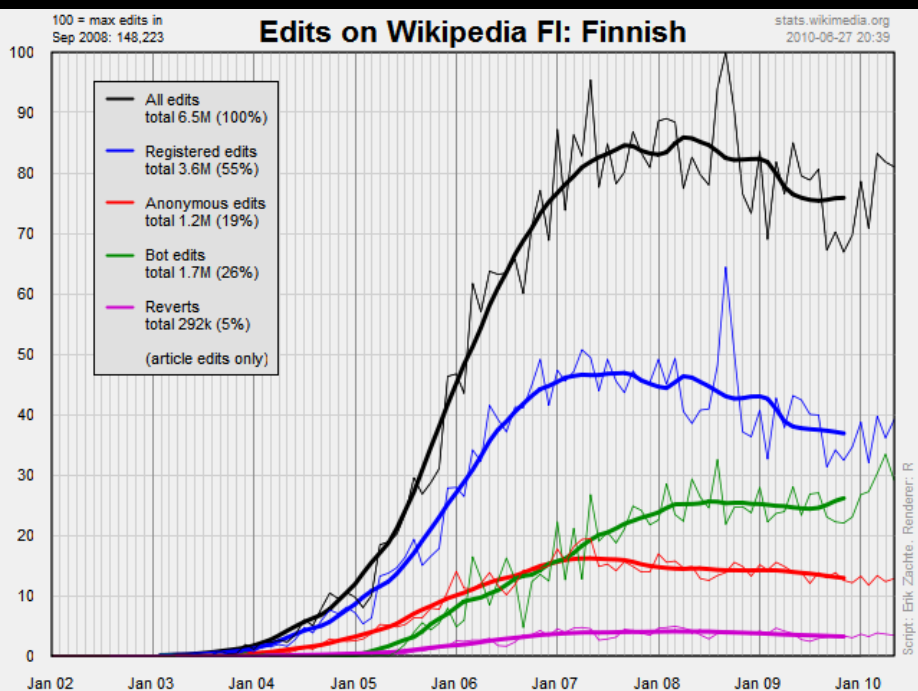
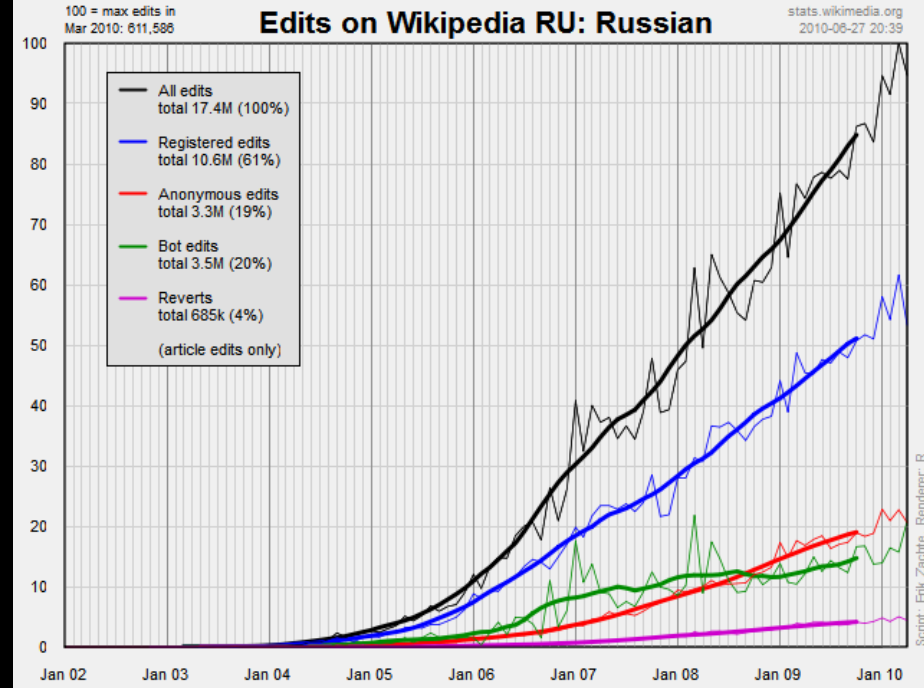
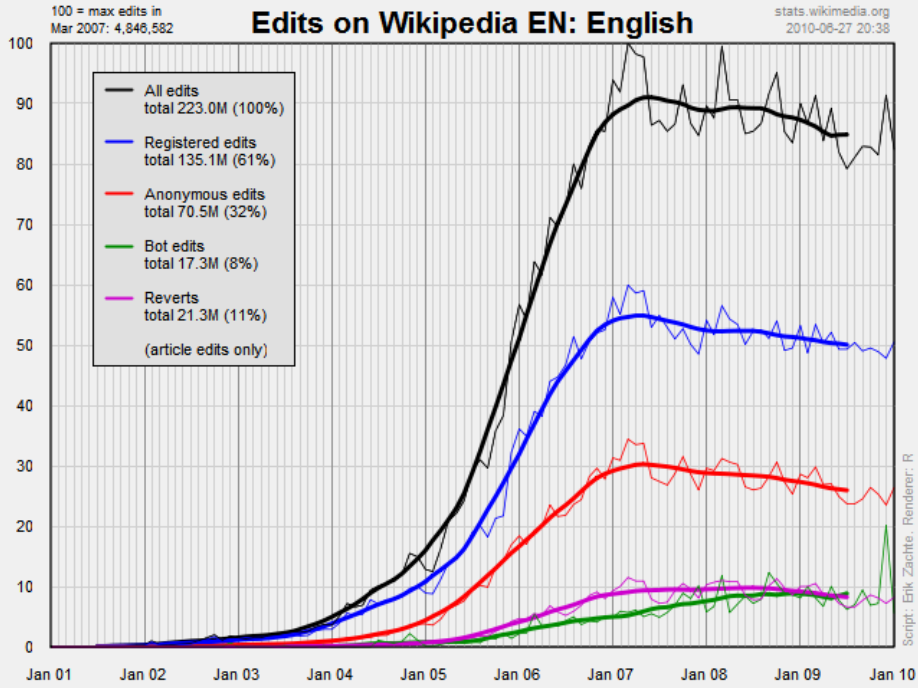
# Edits on all Wikipedias





<http://stats.wikimedia.org/EN/PlotsPngEditHistoryTop.htm>





Reverts





# How to detect reverts

- **Comments**

rv, rvv, vandal, restored, undone, ..(Eng Wp)

- **Checksums (MD5)**




654af3678745236583fc2433g5daaf32

# Calculate MD5 checksum for each revision

- 15A3B...
  - [\(cur | prev\)](#)   18:39, 31 October 2005 [Francs2000](#) ([talk](#) | [contribs](#)) [m](#) (98,186 bytes)  
*(Reverted edits by [82.94.251.206](#) to last version by [Pgk](#))*
- C324B...
  - [\(cur | prev\)](#)   18:38, 31 October 2005 [82.94.251.206](#) ([talk](#)) (45 bytes)
- 15A3B...
  - [\(cur | prev\)](#)   18:38, 31 October 2005 [Pgk](#) ([talk](#) | [contribs](#)) [m](#) (98,186 bytes)  
*(Reverted edits by [82.94.251.206](#) to last version by [Francs2000](#))*
- 9980A...
  - [\(cur | prev\)](#)   18:36, 31 October 2005 [82.94.251.206](#) ([talk](#)) (45 bytes)
- 15A3B...
  - [\(cur | prev\)](#)   18:35, 31 October 2005 [Francs2000](#) ([talk](#) | [contribs](#)) [m](#) (98,186 bytes)  
*(Reverted edits by [68.100.145.172](#) to last version by [Francs2000](#))*
- 67780...
  - [\(cur | prev\)](#)   18:34, 31 October 2005 [68.100.145.172](#) ([talk](#)) (98,271 bytes)
- 15A3B...
  - [\(cur | prev\)](#)   18:15, 31 October 2005 [Francs2000](#) ([talk](#) | [contribs](#)) [m](#) (98,186 bytes)  
*(Reverted edits by [213.20.137.155](#) to last version by [Francs2000](#))*
- 77789...
  - [\(cur | prev\)](#)   18:14, 31 October 2005 [213.20.137.155](#) ([talk](#)) (98,125 bytes)
- 15A3B...
  - [\(cur | prev\)](#)   18:13, 31 October 2005 [Francs2000](#) ([talk](#) | [contribs](#)) [m](#) (98,186 bytes)  
*(Reverted edits by [213.20.137.155](#) to last version by [192.250.112.194](#))*

# Calculate MD5 checksum for each revision



15A3B...	<ul style="list-style-type: none"><li>• (cur   prev) <input type="radio"/> <input type="radio"/> 18:39, 31 October 2005 Francs2000 (talk   contribs) m (98,186 bytes) (Reverted edits by 82.94.251.206 to last version by Pgk)</li></ul>
C324B...	<ul style="list-style-type: none"><li>• (cur   prev) <input type="radio"/> <input type="radio"/> 18:38, 31 October 2005 82.94.251.206 (talk) (45 bytes)</li></ul>
15A3B...	<ul style="list-style-type: none"><li>• (cur   prev) <input type="radio"/> <input type="radio"/> 18:38, 31 October 2005 Pgk (talk   contribs) m (98,186 bytes) (Reverted edits by 82.94.251.206 to last version by Francs2000)</li></ul>
9980A...	<ul style="list-style-type: none"><li>• (cur   prev) <input type="radio"/> <input type="radio"/> 18:36, 31 October 2005 82.94.251.206 (talk) (45 bytes)</li></ul>
15A3B...	<ul style="list-style-type: none"><li>• (cur   prev) <input type="radio"/> <input type="radio"/> 18:35, 31 October 2005 Francs2000 (talk   contribs) m (98,186 bytes) (Reverted edits by 68.100.145.172 to last version by Francs2000)</li></ul>
67780...	<ul style="list-style-type: none"><li>• (cur   prev) <input type="radio"/> <input type="radio"/> 18:34, 31 October 2005 68.100.145.172 (talk) (98,271 bytes)</li></ul>
15A3B...	<ul style="list-style-type: none"><li>• (cur   prev) <input type="radio"/> <input type="radio"/> 18:15, 31 October 2005 Francs2000 (talk   contribs) m (98,186 bytes) (Reverted edits by 213.20.137.155 to last version by Francs2000)</li></ul>
77789...	<ul style="list-style-type: none"><li>• (cur   prev) <input type="radio"/> <input type="radio"/> 18:14, 31 October 2005 213.20.137.155 (talk) (98,125 bytes)</li></ul>
15A3B...	<ul style="list-style-type: none"><li>• (cur   prev) <input type="radio"/> <input type="radio"/> 18:13, 31 October 2005 Francs2000 (talk   contribs) m (98,186 bytes) (Reverted edits by 213.20.137.155 to last version by 192.250.112.194)</li></ul>

# How to detect reverts

- Comments

  - Language dependent

    - rv, rvv, vandal, restored, undone, ..

  - More false positives

    - e.g. “not vandalism, see .....

- MD5 checksums

  - Language independent

  - Few false positives

    - some status pages

# Rankings

English Wikipedia  
dump January 2010

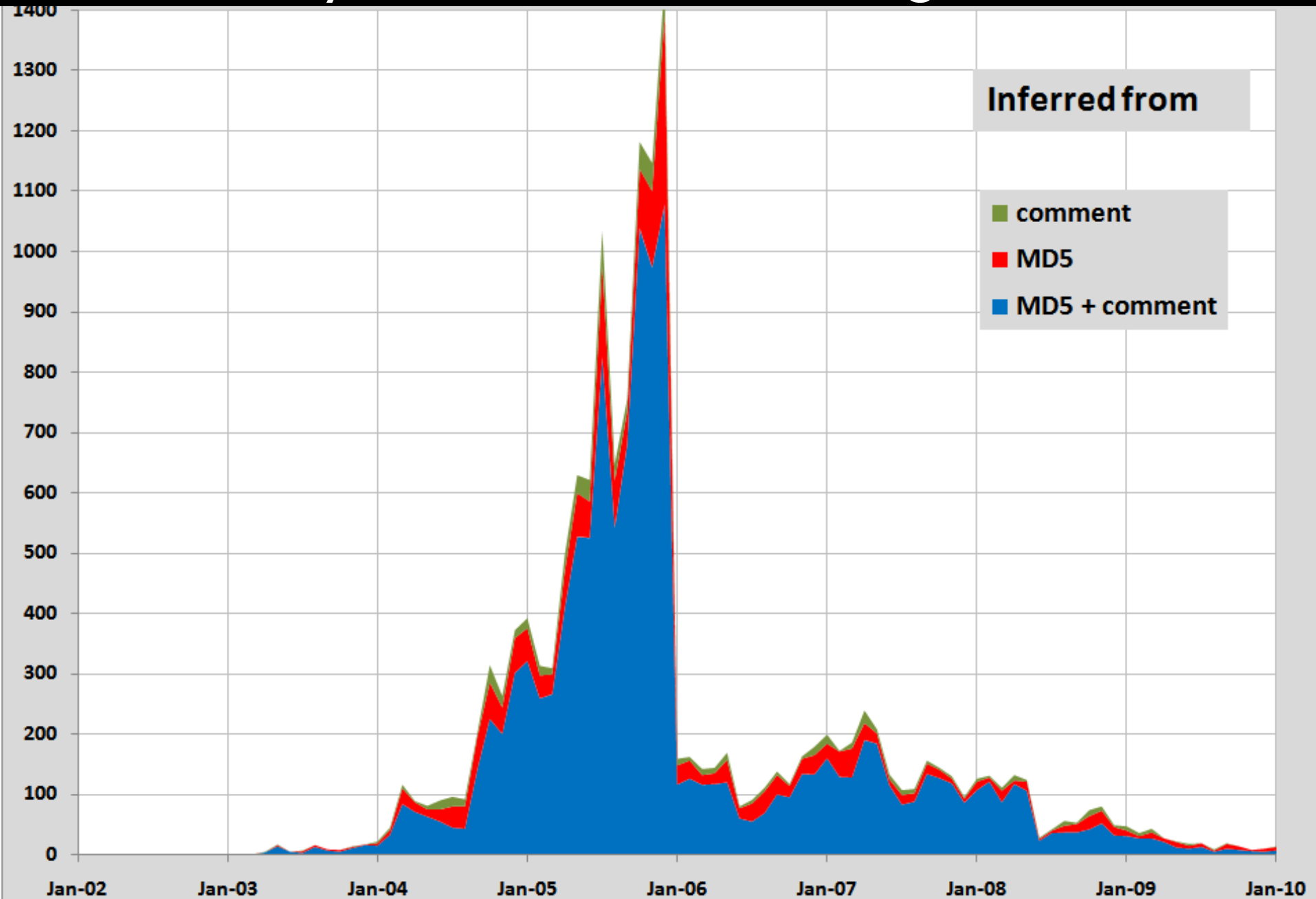
Most Active Reverters		Most Reverted Editors	
# Rv	User/Bot	# Rv	User/Bot
642178	<a href="#">ClueBot</a>	36199	<a href="#">ClueBot</a>
141294	<a href="#">AntiVandalBot</a>	30637	<a href="#">RussBot</a>
134601	<a href="#">J.delanoy</a>	16360	<a href="#">Redirect fixer</a>
134258	<a href="#">VoABot II</a>	14467	<a href="#">Xqbot</a>
103190	<a href="#">XLinkBot</a>	11817	<a href="#">DarknessBot</a>
101804	<a href="#">Alansohn</a>	11113	<a href="#">AntiVandalBot</a>
100659	<a href="#">MartinBot</a>	8957	<a href="#">OrphanBot</a>
69898	<a href="#">Tawkerbot2</a>	8134	<a href="#">SmackBot</a>
68355	<a href="#">Epbr123</a>	7726	<a href="#">XLinkBot</a>
64604	<a href="#">Discospinster</a>	7526	<a href="#">VoABot II</a>
64367	<a href="#">Can't sleep%2C clown will eat me</a>	7455	<a href="#">MartinBot</a>
60654	<a href="#">Bobo192</a>	6622	<a href="#">Cydebot</a>
52111	<a href="#">The Thing That Should Not Be</a>	5953	<a href="#">J.delanoy</a>
49949	<a href="#">S. J. Liebowitz</a>	5204	<a href="#">S. J. Liebowitz</a>

# Rankings

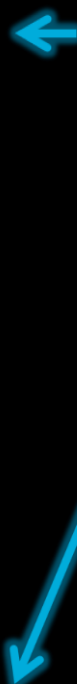
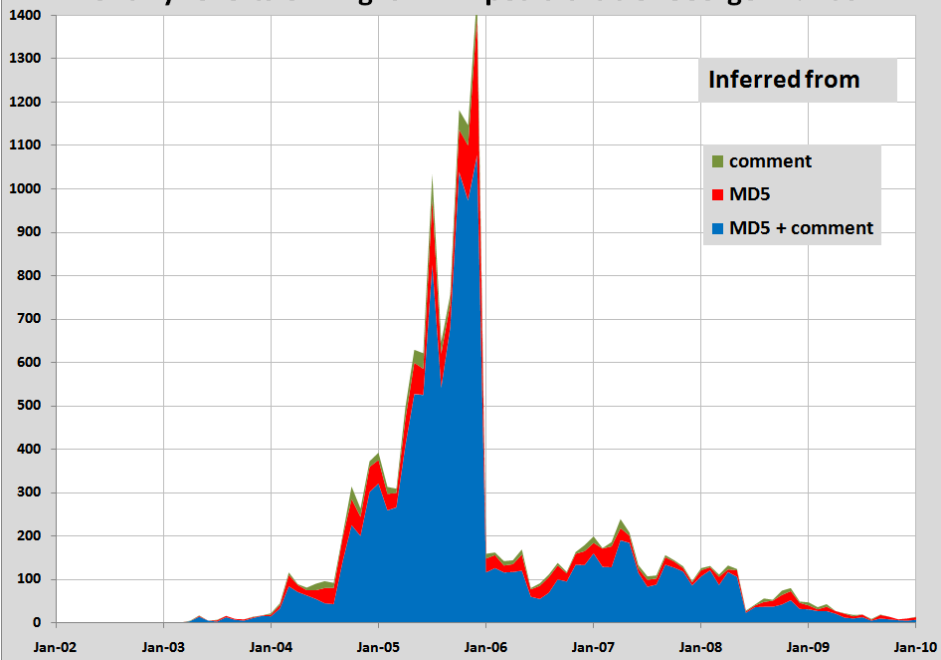
English Wikipedia  
dump January 2010

Most Reverted Articles	
# Rv	Article
14972	<a href="#">George W. Bush</a>
11357	<a href="#">Wikipedia</a>
6527	<a href="#">United States</a>
6103	<a href="#">Adolf Hitler</a>
5211	<a href="#">Jesus</a>
4905	<a href="#">Global warming</a>
4351	<a href="#">Michael Jackson</a>
4310	<a href="#">World War II</a>
4300	<a href="#">Barack Obama</a>
4098	<a href="#">Wii</a>
4064	<a href="#">RuneScape</a>
4059	<a href="#">Muhammad</a>
3998	<a href="#">Britney Spears</a>
3911	<a href="#">Lionel Messi</a>

# Monthly reverts on WP:EN 'George W. Bush'



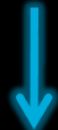
Monthly reverts on English Wikipedia article 'George W. Bush'



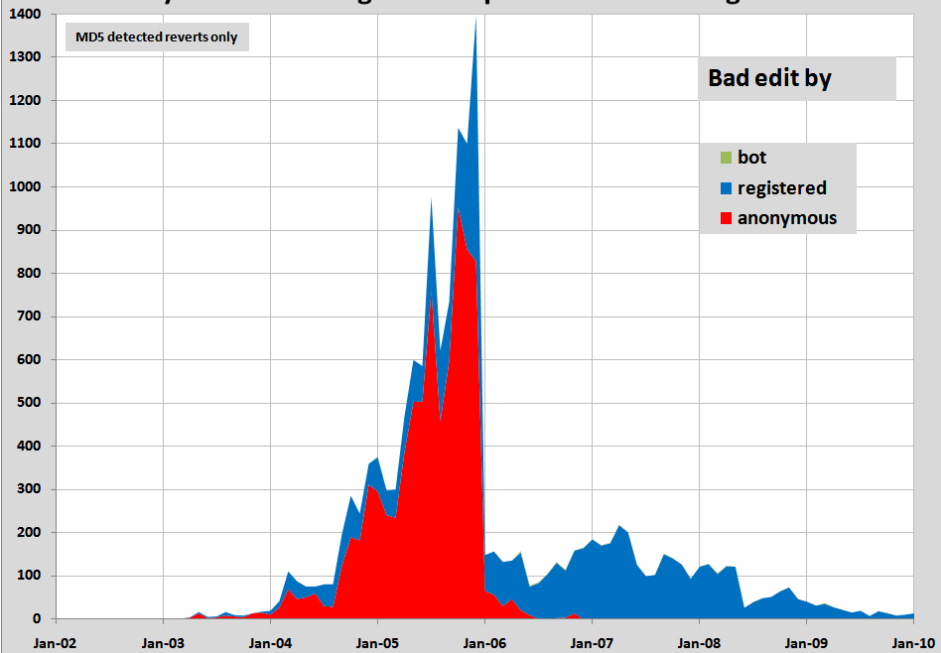
← Monthly reverts

Reverted editors

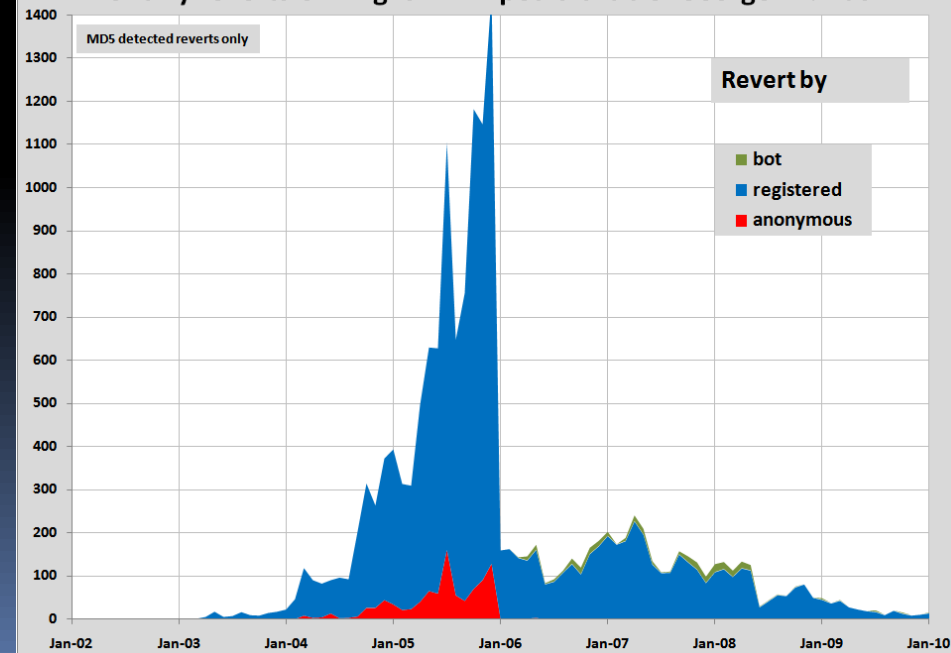
Reverting editors



Monthly reverts on English Wikipedia article 'George W. Bush'

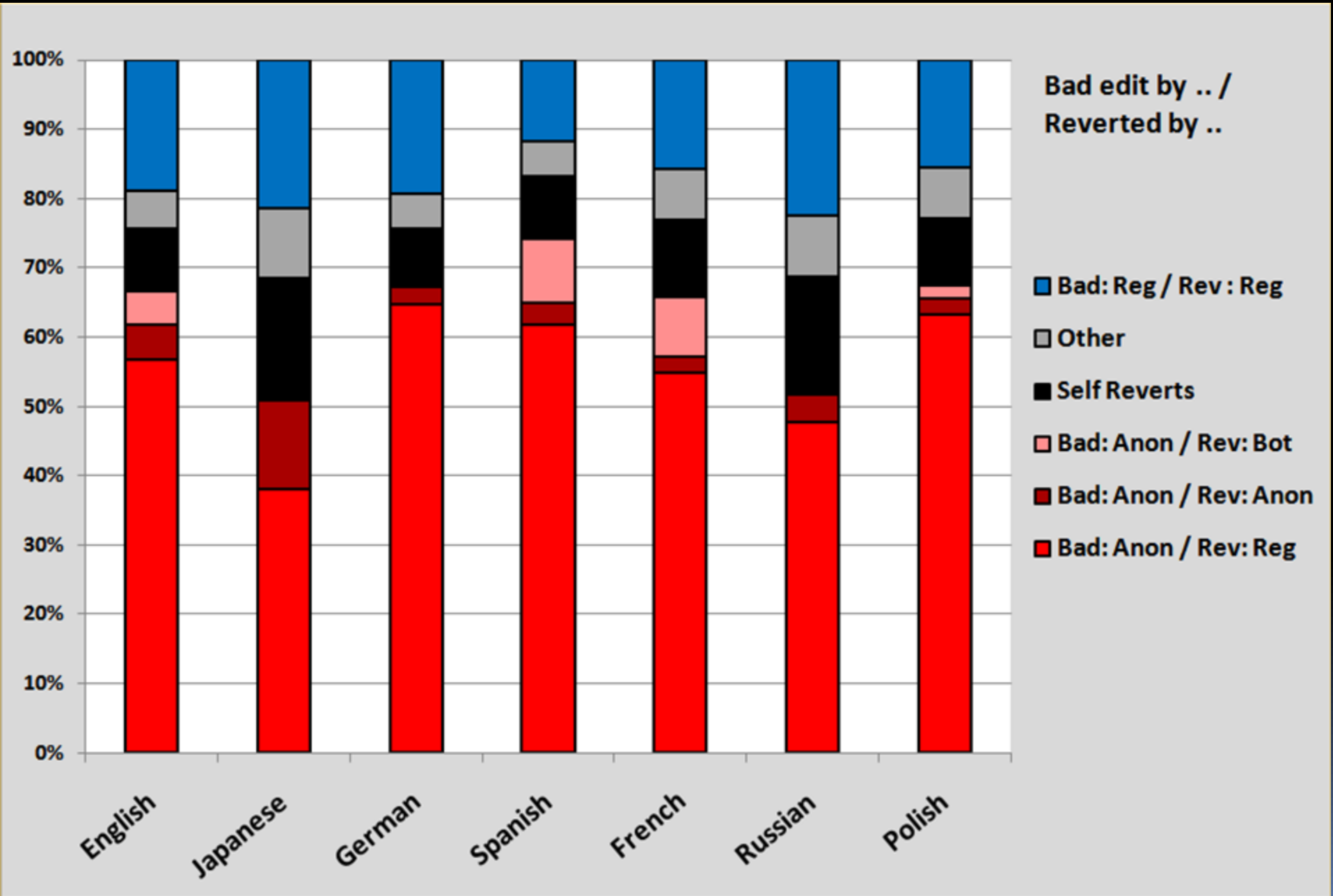


Monthly reverts on English Wikipedia article 'George W. Bush'

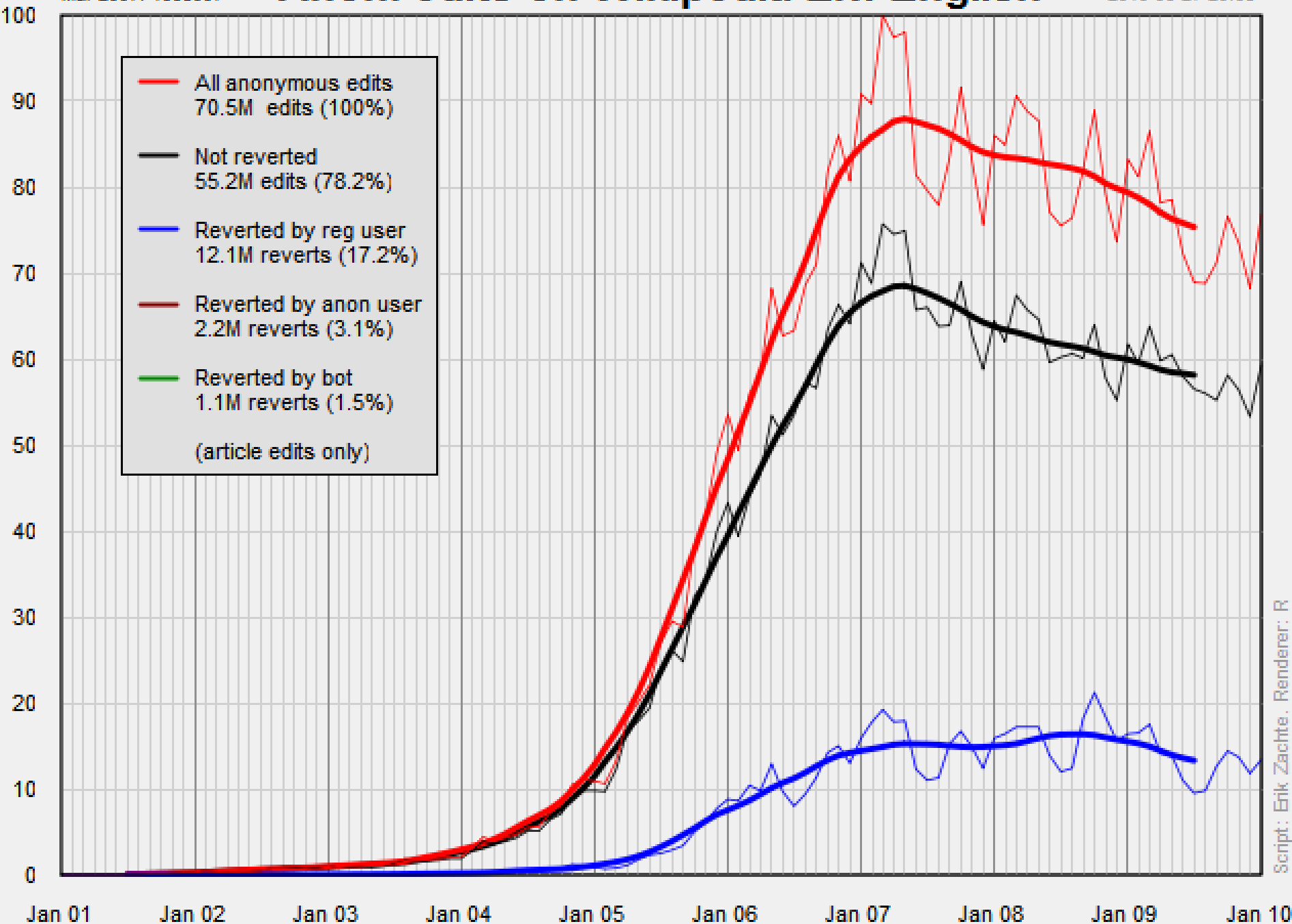




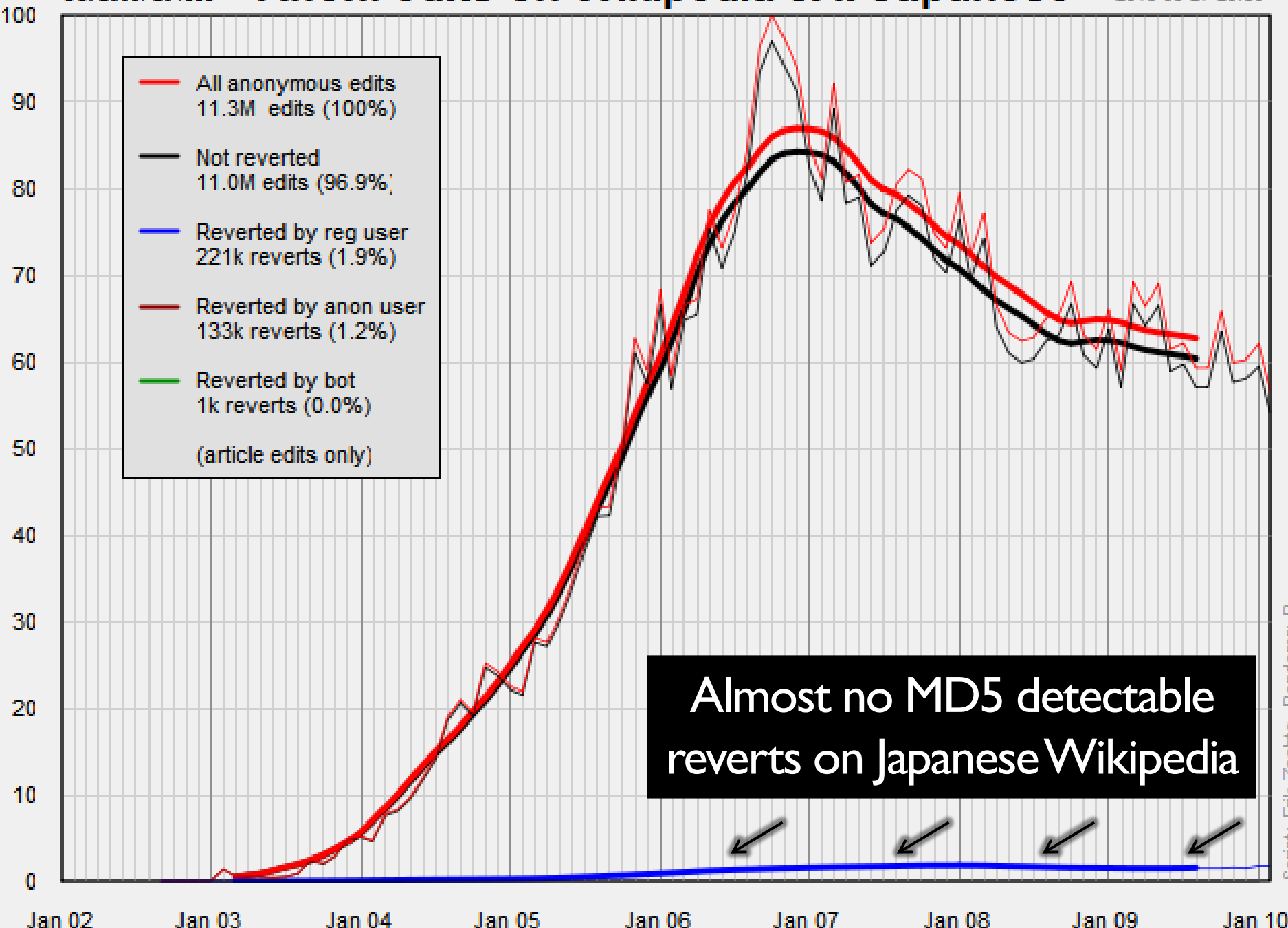
# Breakdown of reverts on largest Wikipedias



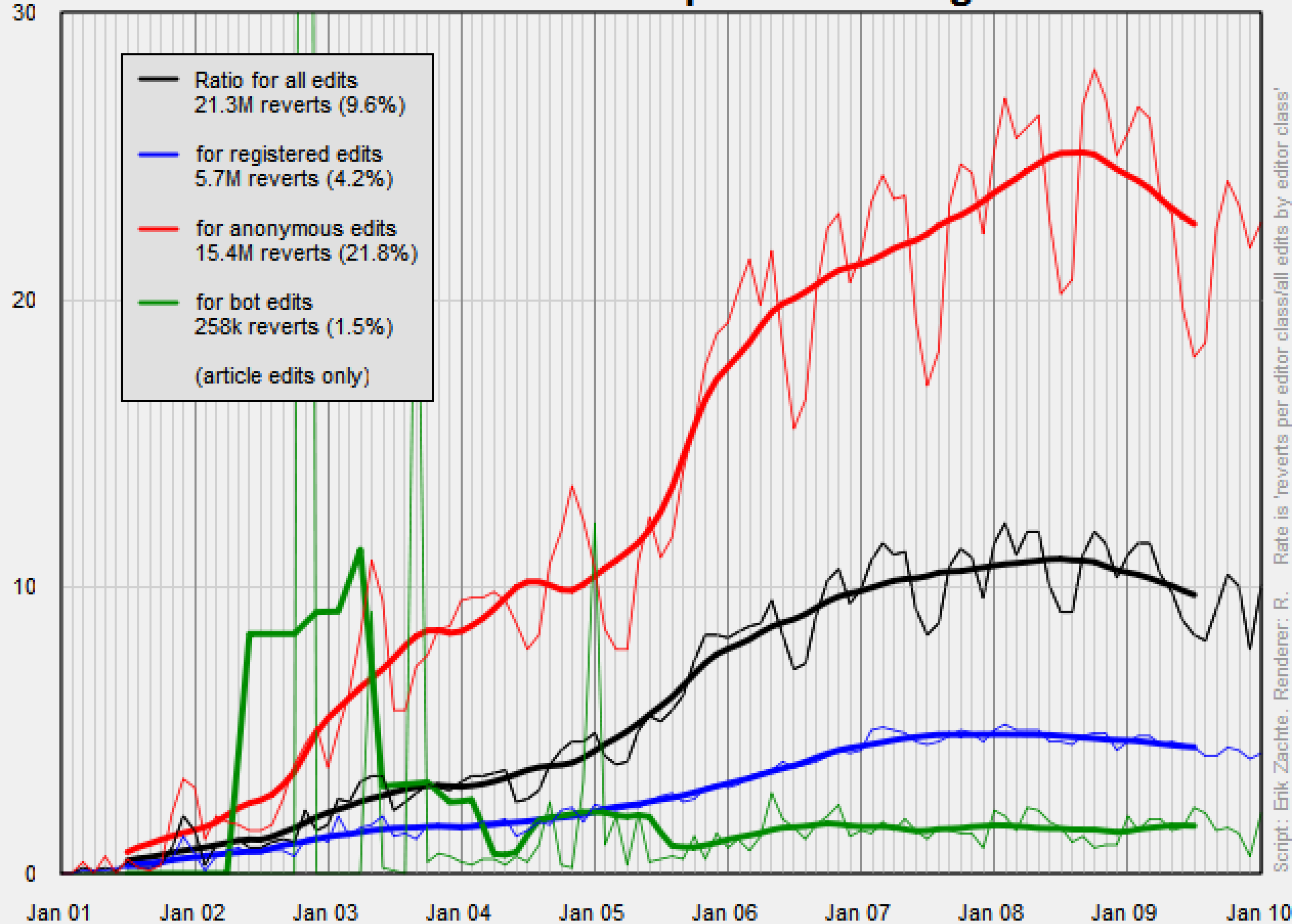
# Anon. edits on Wikipedia EN: English



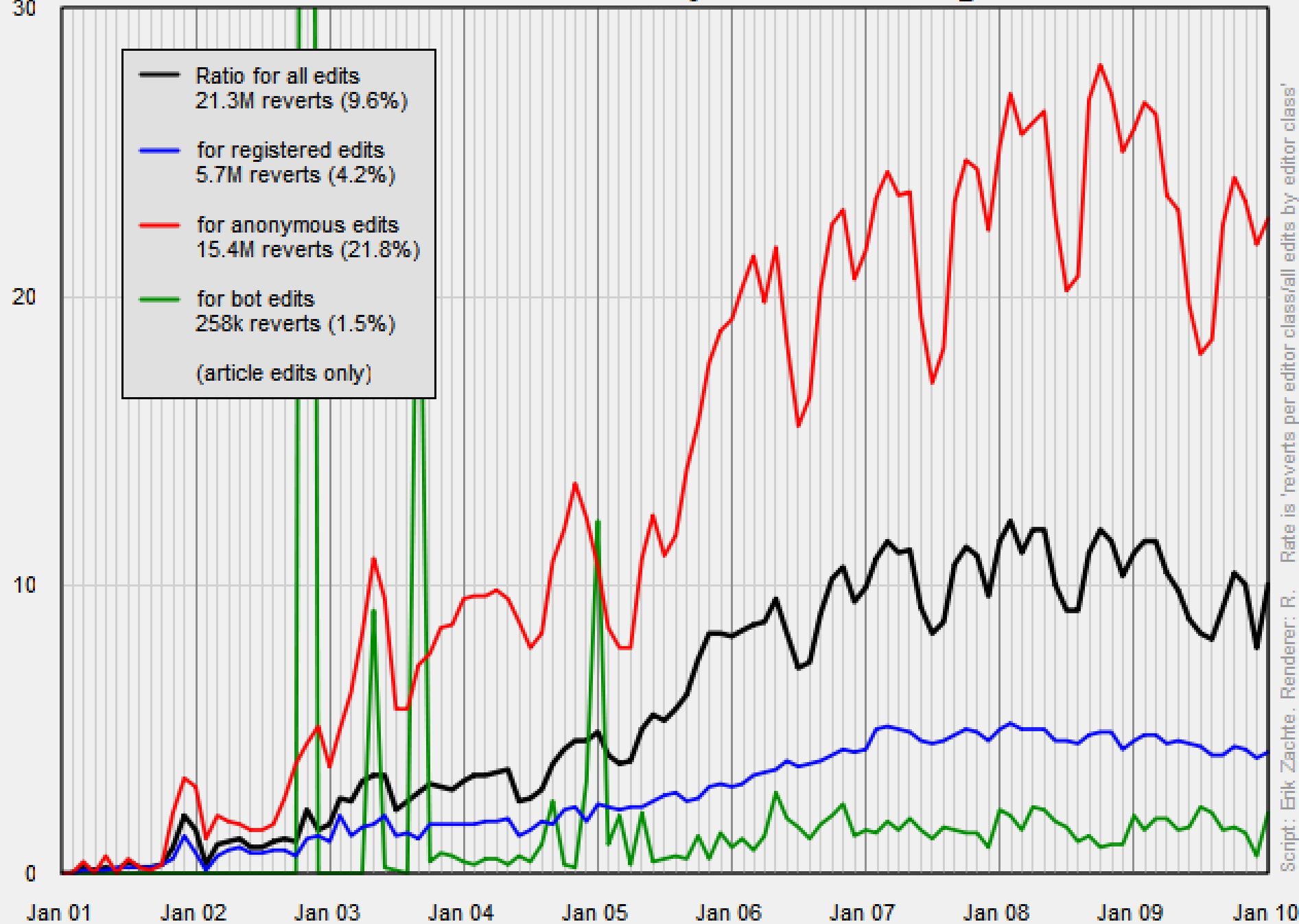
# Anon. edits on Wikipedia JA: Japanese

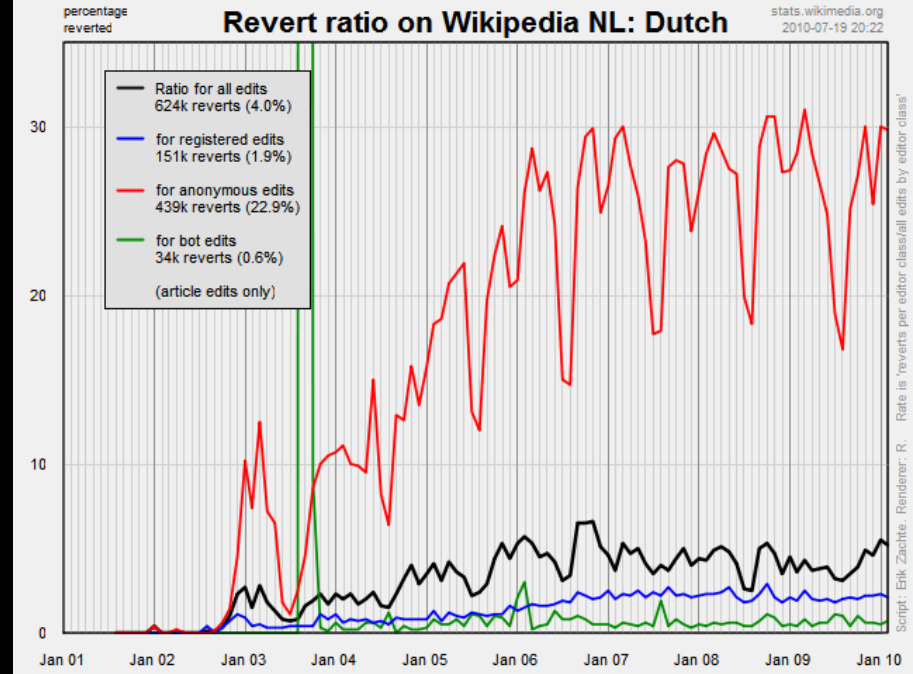
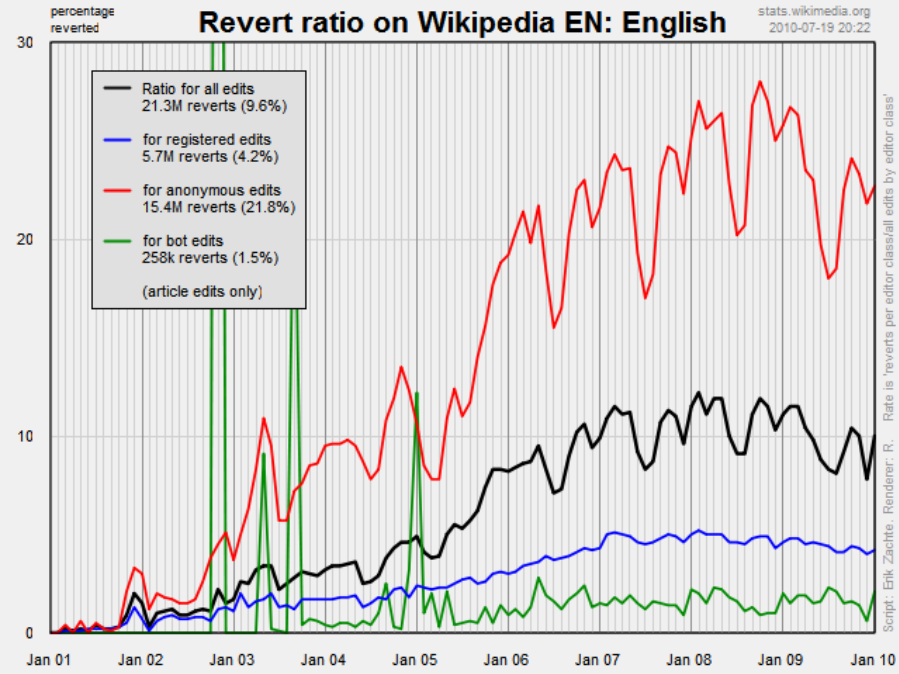


# Revert ratio on Wikipedia EN: English

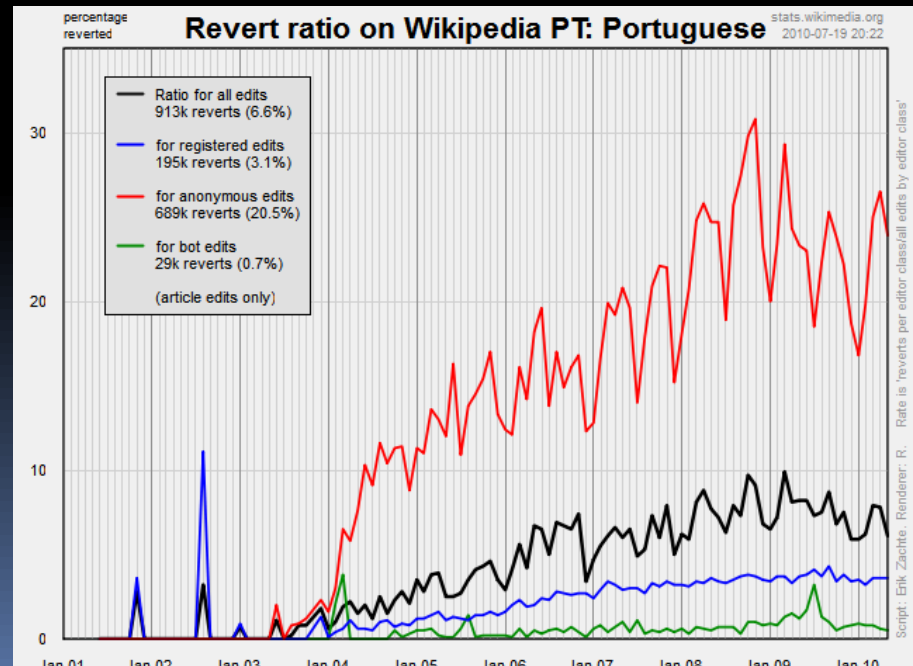
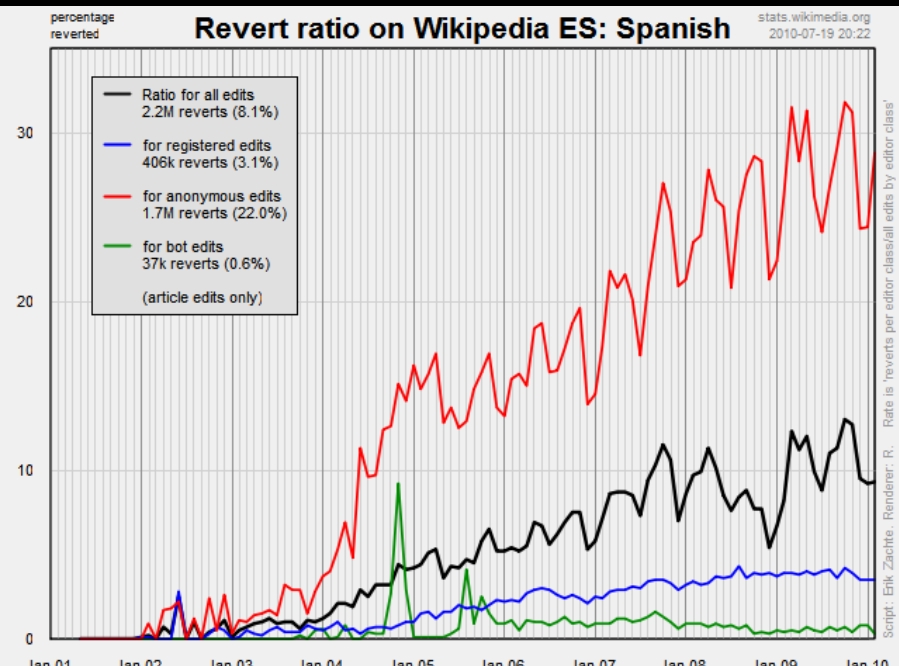


# Revert ratio on Wikipedia EN: English





**Strong seasonal pattern for anon reverts**





# Flagged Revisions on German Wikipedia

# Flagged Revisions / Pending Changes

- **Is it effective?**
- **Is it destructive?**
- **On what scale is it being used?**



# Flagged revisions / Pending Changes

- **Is it effective?**

Does it shield bad contributions from the general public?

**Does it discourage bad contributions ?**

Are good contributions released timely to general public?

- **Is it destructive?**

**Does it break or discourage collaborative processes?**

- **On what scale is it being used? (+ trend)**

To how many articles does this protection apply

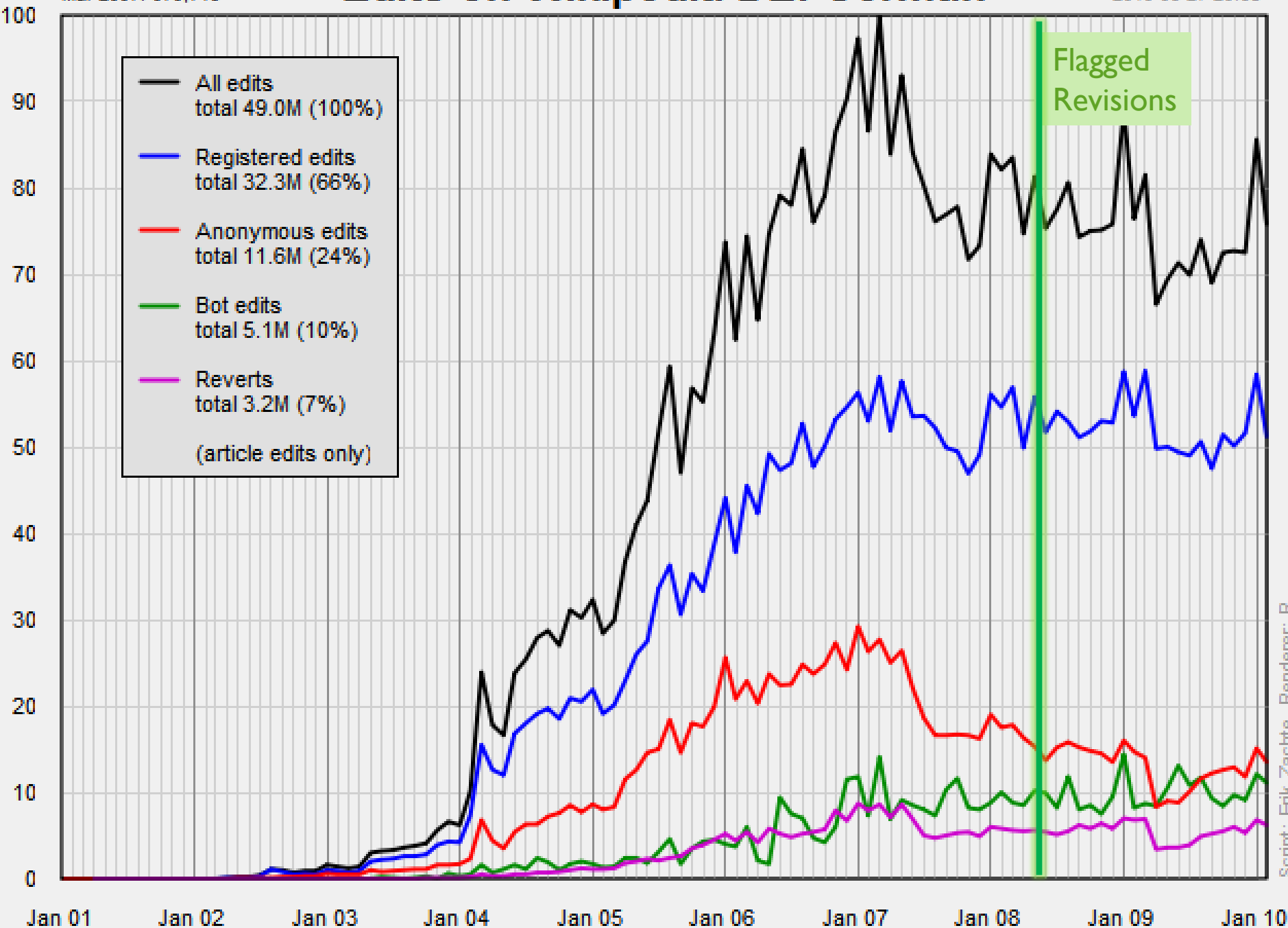
Can the community absorb the amount of work ?

# Flagged Revisions

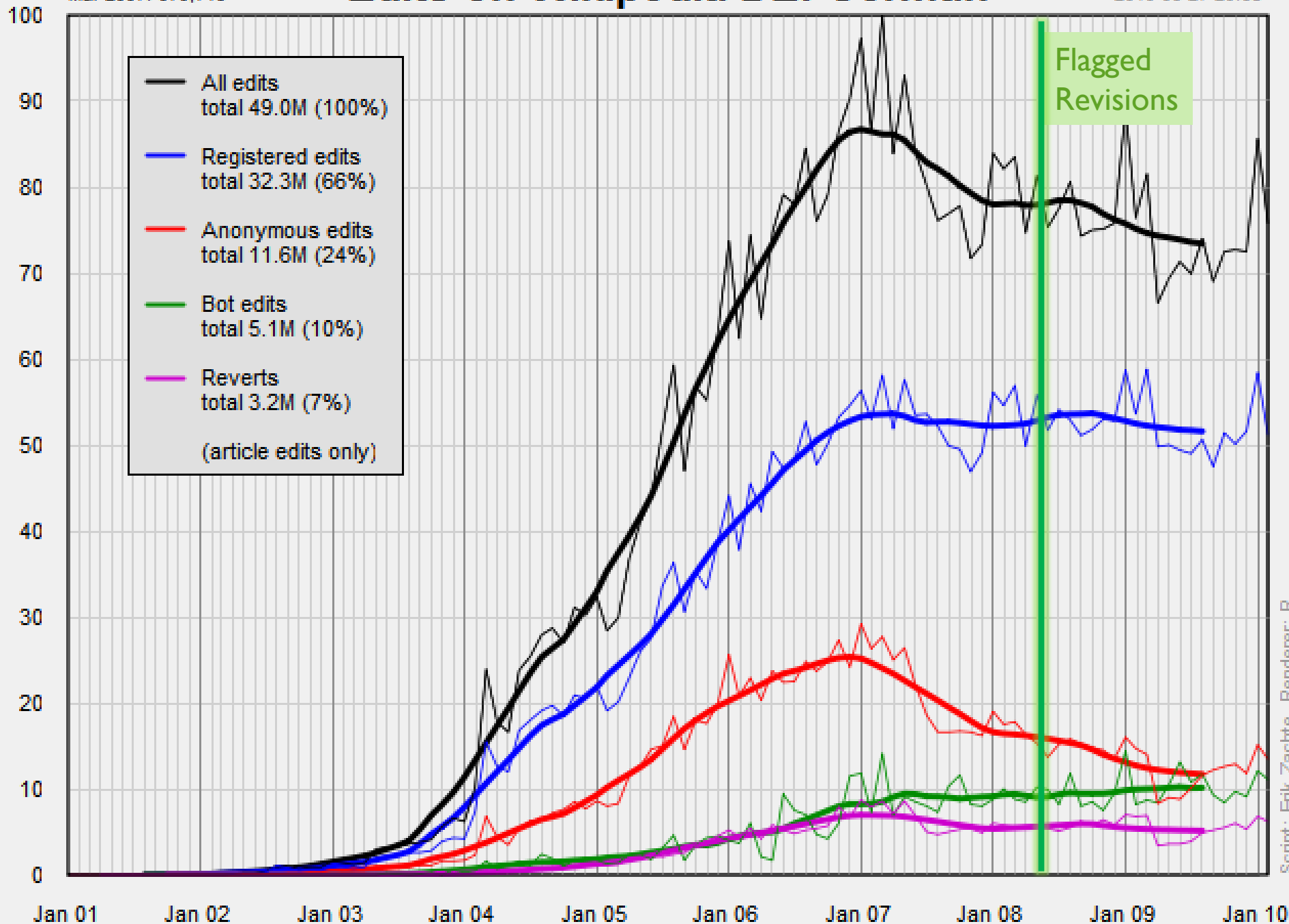
introduced May 6, 2008

on German Wikipedia

# Edits on Wikipedia DE: German

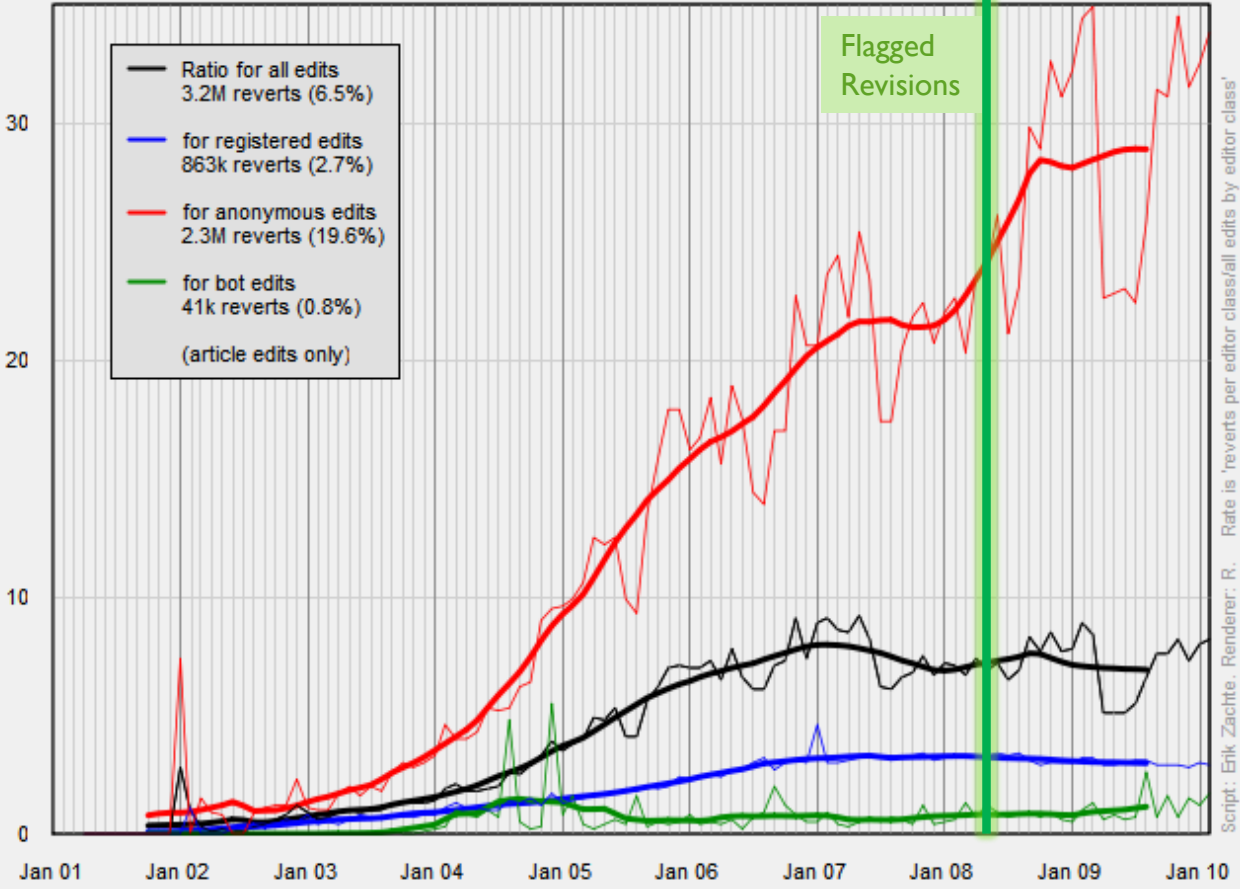


# Edits on Wikipedia DE: German



# Revert ratio on Wikipedia DE: German

## Revert ratio climbs

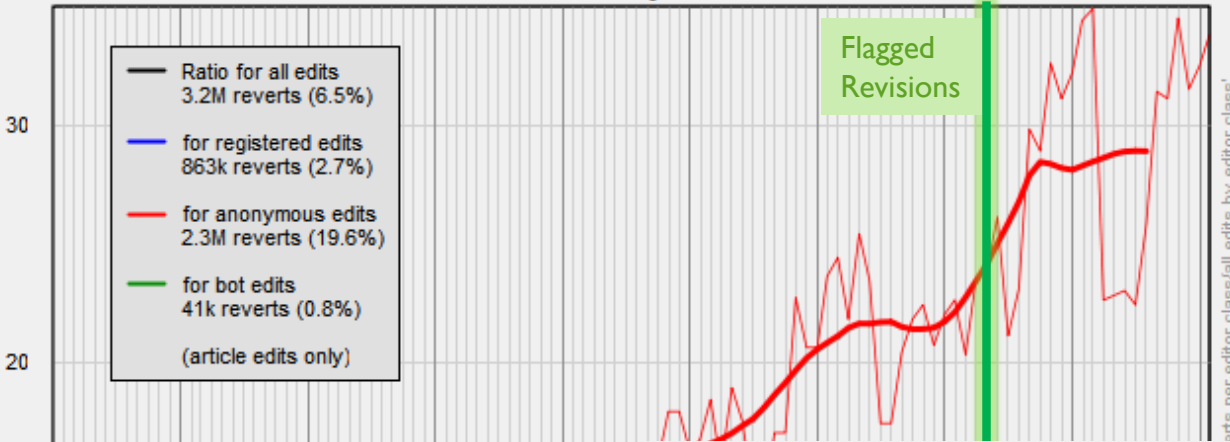


percentage  
reverted

# Revert ratio on Wikipedia DE: German

stats.wikimedia.org  
2010-07-19 20:22

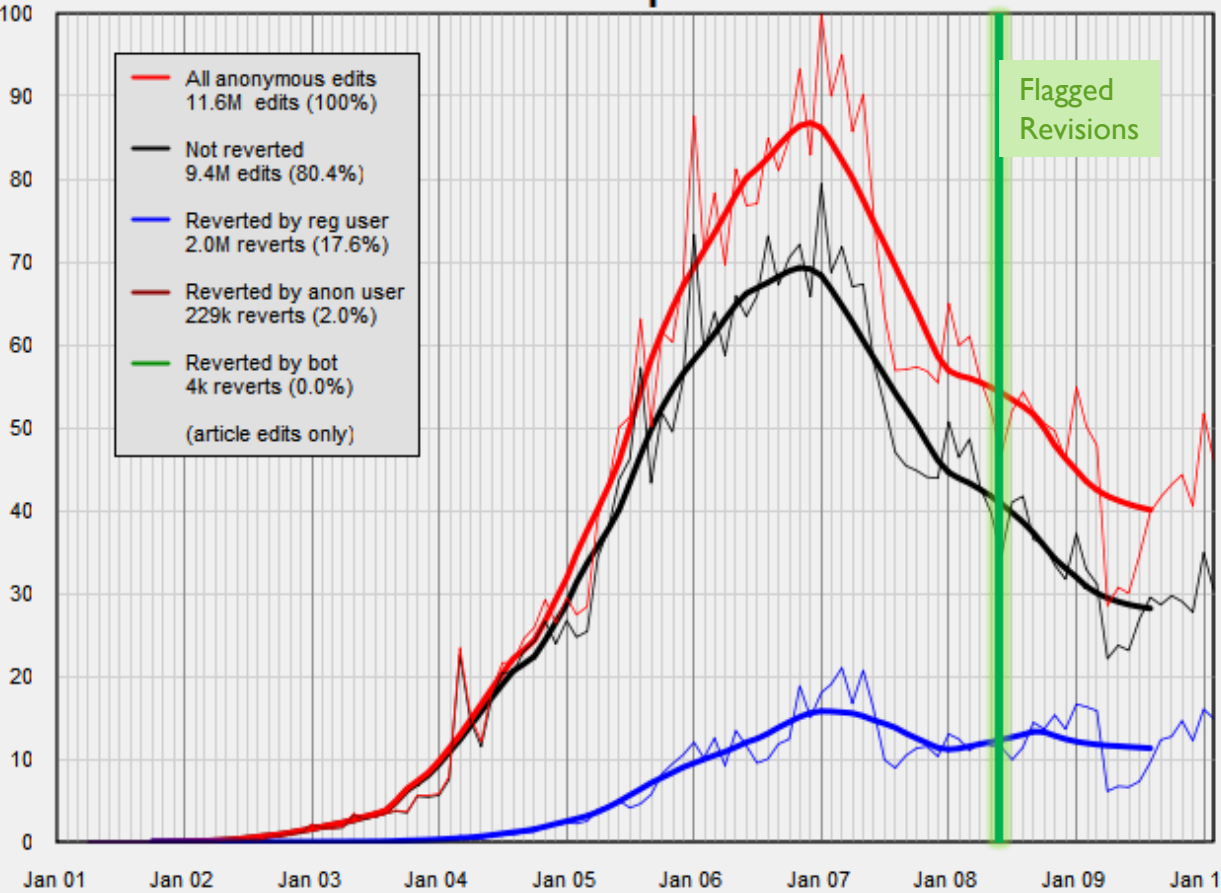
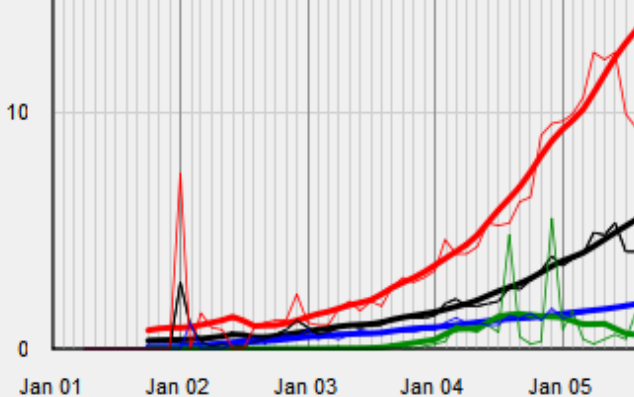
# Revert ratio climbs



100 = max anon edits in  
Jan 2007: 288202

# Anon. edits on Wikipedia DE: German

stats.wikimedia.org  
2010-06-27 20:38



Less anon edits  
Revert volume stable

Script: Erik Zachite, Renderer: R

# Flagged Revisions on German Wikipedia

Edit-Revert charts suggest

Does it effect registered editors? **No**  
(like expected)

Does it discourage vandals? **No**

Does it discourage good anon editors?  
**Seems not** (no sudden trend changes)

# Flagged Revisions / Pending Changes on all Wikipedias

Follow up with same data files

Does it shield bad content from public?

Are good contributions released timely?



# Questions ?



Images from commons:

- [http://commons.wikimedia.org/wiki/File:Bundesarchiv\\_Bild\\_183-27924-0001,\\_MTS\\_Magdeburg-S%C3%BCdwest,\\_Setzen\\_einer\\_Dorfzeitung.jpg](http://commons.wikimedia.org/wiki/File:Bundesarchiv_Bild_183-27924-0001,_MTS_Magdeburg-S%C3%BCdwest,_Setzen_einer_Dorfzeitung.jpg)
- [http://commons.wikimedia.org/wiki/File:Konrad\\_Plautz,\\_Schiedsrichter\\_%284%29.jpg](http://commons.wikimedia.org/wiki/File:Konrad_Plautz,_Schiedsrichter_%284%29.jpg)
- [http://commons.wikimedia.org/wiki/File:Canberra\\_-\\_Maltese\\_Flags.jpg](http://commons.wikimedia.org/wiki/File:Canberra_-_Maltese_Flags.jpg)
- [http://commons.wikimedia.org/wiki/File:Wikimania\\_2009\\_-\\_Questions.jpg](http://commons.wikimedia.org/wiki/File:Wikimania_2009_-_Questions.jpg)

Extra slides not shown at Wikimania

# Wikistats: Articles ranked by total edit count

## Wikimedia articles in any project ranked by total edit count

[Home](#)

See also [Wikimedia articles in any project ranked by uncompressed archive size](#)

Rank	Total Size	Edits	Avg bytes/ Submit	Lang	Project	Article
1	1.5 Gb	642604	2500	en	wikipedia	<a href="#">Wikipedia:Administrator intervention against vandalism</a>
2	117 Gb	419173	299529	en	wikipedia	<a href="#">Wikipedia:Administrators' noticeboard/Incidents</a>
3	1.1 Gb	326332	3676	en	wikipedia	<a href="#">Wikipedia:Sandbox/Archive</a>
4	21 Gb	290652	76102	meta	wikipedia	<a href="#">User:COIBot/LinkReports</a>
5	887 Mb	257435	3612	en	wikipedia	<a href="#">Wikipedia:Sandbox</a>
6	1.1 Gb	226430	5081	en	wikipedia	<a href="#">User:Cyde/List of candidates for speedy deletion/Subpage</a>
7	971 Mb	217245	4686	it	wikinews	<a href="#">Template:Index</a>
8	38 Gb	204223	201046	en	wikipedia	<a href="#">Wikipedia:Reference desk/Science</a>
9	11 Gb	191582	63893	en	wikipedia	<a href="#">Wikipedia:WikiProject Spam/LinkReports</a>
10	25 Gb	186590	146493	en	wikipedia	<a href="#">Wikipedia:Reference desk/Miscellaneous</a>
11	27 Gb	184235	158032	en	wikipedia	<a href="#">Template talk:Did you know</a>
12	22 Gb	169831	137117	en	wikipedia	<a href="#">Wikipedia:Help desk</a>
13	4.0 Gb	164436	26054	en	wikipedia	<a href="#">Wikipedia:Requests for page protection</a>
14	28 Gb	155629	193435	en	wikipedia	<a href="#">Wikipedia:Administrators' noticeboard</a>
15	20 Gb	148865	146221	en	wikipedia	<a href="#">Wikipedia:Reference desk/Humanities</a>
16	699 Mb	142395	5148	en	wikipedia	<a href="#">Wikipedia:Introduction</a>
17	310 Mb	123317	2638	it	wikinews	<a href="#">Template:Currency</a>
18	176 Mb	108792	1697	es	wikipedia	<a href="#">Wikipedia:Zona de pruebas</a>
19	22 Gb	103150	227595	nl	wikipedia	<a href="#">Wikipedia:De kroeg</a>
20	210 Mb	92448	2387	en	wikipedia	<a href="#">Wikipedia:Usernames for administrator attention</a>

# Wikistats: Articles ranked by total edit count

## Wikimedia articles in any project ranked by total edit count

[Home](#)

See also [Wikimedia articles in any project ranked by total edit count](#)

[intervention against vandalism](#)

Rank	Total Size	Edits	Avg bytes/ Submit	Lang	Project	Article
1	1.5 Gb	642604	2500	en	wikipedia	<a href="#">Wikipedia:Administrators' intervention against vandalism</a>
2	117 Gb	419173	299529	en	wikipedia	<a href="#">Wikipedia:Administrators' noticeboard/Incidents</a>
3	1.1 Gb	326332	3676	en	wikipedia	<a href="#">Wikipedia:Sandbox/Archive</a>
4	21 Gb	290652	76102	meta	wikipedia	<a href="#">User:COIBot/LinkReports</a>
5	887 Mb	257435	3612	en	wikipedia	<a href="#">Wikipedia:Sandbox</a>
6	1.1 Gb	226430	5081	en	wikipedia	<a href="#">User:Cydel/ist of candidates for speedy deletion/Subpage</a>

Total Size	Edits	Avg bytes/ Submit	Lang
1.5 Gb	642604	2500	en
117 Gb	419173	299529	en
1.1 Gb	326332	3676	en
21 Gb	290652	76102	meta

[attention](#)

# Breakdowns

Reverted editor		Revert by		Self revert	
Reg user	26.8%	Reg user	22.3%	Self	3.4%
				Others	18.9%
		Anon user	3.5%		
		Bot	1.0%		
Anon user	72.0%	Reg user	56.8%		
		Anon user	10.2%	Self	5.3%
				Others	4.9%
		Bot	5.0%		
Bot	1.2%	Reg user	0.5%		
		Anon user	0.3%		
		Bot	0.3%	Self	0.2%
				Others	0.2%

Reg = registered, Anon = anonymous

but percentages are  
article namespace only

	2005	2004	2003	2002	2001
9%	6.1%	7.5%	11.6%	10.4%	
0%	25.1%	30.4%	25.9%	27.3%	
8%	4.3%	4.7%	6.9%	8.5%	
3%	0.3%	0.0%	-	-	
0%	51.7%	45.7%	38.4%	38.8%	
5%	7.7%	8.5%	10.5%	9.6%	
9%	3.1%	2.4%	6.3%	3.4%	
0%	0.1%	0.2%	0.1%	1.3%	
4%	0.9%	0.2%	0.3%	0.6%	
0%	0.0%	0.0%	0.0%	-	
1%	0.8%	0.4%	-	-	
0%	0.0%	0.0%	-	-	

totals reverts per year

419 k	4.8 M	5.7 M	5.5 M	3.7 M	978 k	201 k	36 k	6 k	469
-------	-------	-------	-------	-------	-------	-------	------	-----	-----