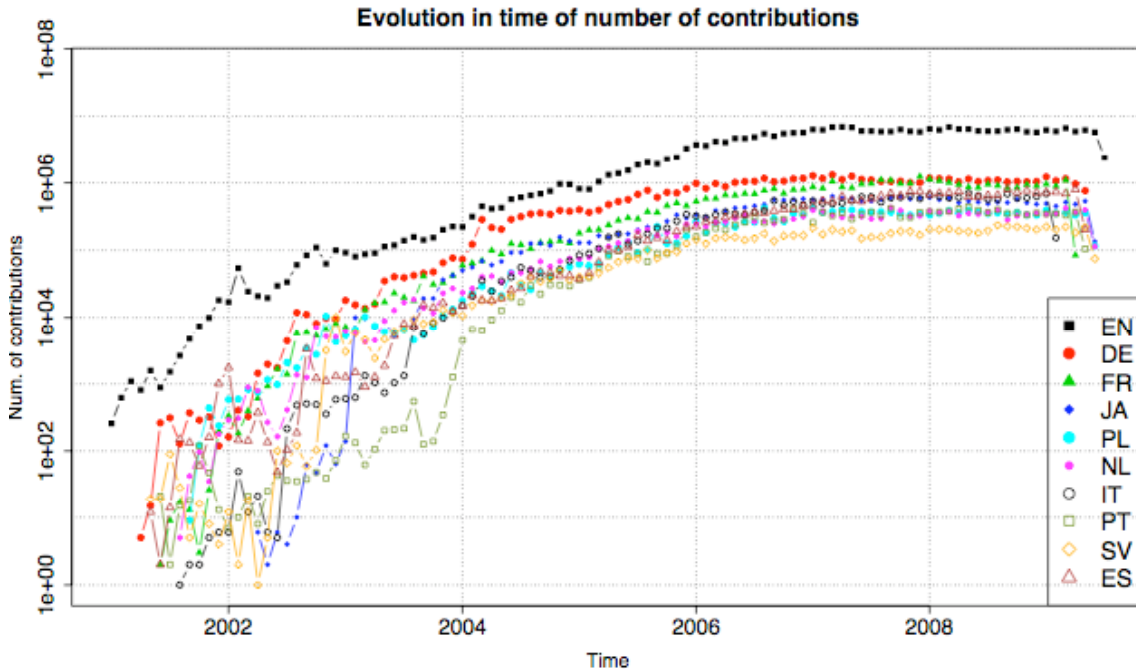


Interview with José Felipe Ortega Soto (PhD who analyzed Wikipedia's evolution)

19 November, 2009

What are your thoughts on the stagnation of Wikipedia's growth? Based on quantitative data, can we conclude that Wikipedia has stagnated?



Be careful with the wording. "Stagnate" implicitly means something wrong in happening. This doesn't have to be true. We are just watching a stabilization. We don't know yet if this is due to internal/external factors, because the community has managed to reach some upper limit, or a combination of both.

Steady-state, stabilize, plateau, constant rate, are preferable in this case.

It's like an electric circuit providing constant power. At first, you have the transitory state, and a quick raise up to the working level. Then, the voltage/current becomes stable. This is not wrong, but just a working condition. The challenge here is that we don't know what to expect from the "Wikipedia circuit" beforehand, and thus we're just trying to perform some reverse engineering.

If we were to enter a decreasing slope period, *that* would be worrying. But we're not there, at least not yet.

What effect will this “stabilization” have on Wikipedia?

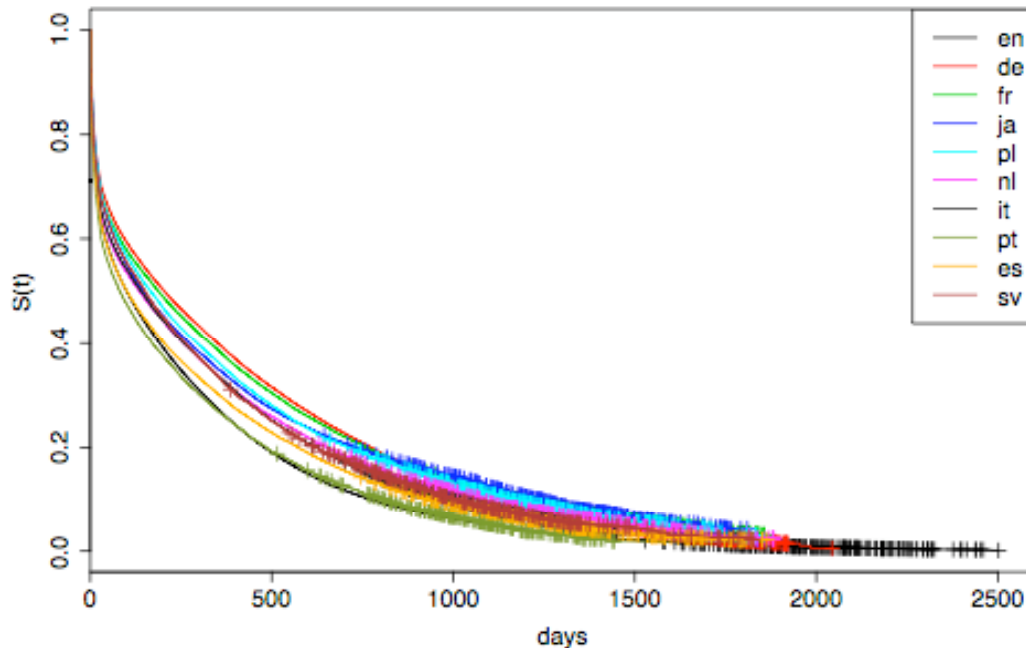
Mean and median of number of distinct logged authors in FAs and non-FAs.

Lang.	Type	Mean #logged auth.	Median #logged auth.	Lang.	Type	Mean #logged auth.	Median #logged auth.
EN	FAs	216.4	113	IT	FAs	50.35	37
	non-FAs	13.97	6		non-FAs	5.491	3
DE	FAs	80.17	57	NL	FAs	46.03	37
	non-FAs	11.57	7		non-FAs	6.301	4
FR	FAs	58.27	41	PT	FAs	37.75	28
	non-FAs	7.493	4		non-FAs	4.242	2
PL	FAs	38.62	27	ES	FAs	49.99	32
	non-FAs	5.473	3		non-FAs	6.828	4
JA	FAs	60.65	49.5	SV	FAs	30.4	23
	non-FAs	7.661	4		non-FAs	4.654	3

We have shown that quality (i.e., getting the FA distinction) is highly correlated with number of different authors and number of revisions (so far, necessary but not sufficient condition). Thus, with a constant rate of edits I expect articles to improve more slowly. Further investigation should prove or refute this hypothesis. The rapid net loss of editors is more worrying. Eventually, community may lack of enough effort to maintain the wiki in proper conditions.

Several editors and experts think that policy proliferation has harmed community growth. Based on your analysis of the different language Wikipedias and their policies, can you draw any conclusions about the effect of policy on editor commitment?

Survival time for logged authors



I've tried to compare the results among different languages to check that question. Surprisingly, the answer is no: there isn't any influence of those policies on the evolution of activity patterns or the commitment of editors in different communities.

The only one difference is the level of monthly number of edits and active editors reached by each language. This is determined only by the initial growing rate (e.g. German Wikipedia received more edits at first than Dutch Wikipedia, and that's why it stabilized at a higher level).

What still puzzles the research community is why a project fed by spontaneous contributions from editors consistently reaches a self-sustained working condition (the monthly Gini coefficient reach a constant level, at the same values), disregarding the language version you measure. This is the cause of eventually entering a "plateau" stage in all versions.

For sure, I'm confident to say that EN, DE, FR and ES all have strong policies and codes of ethics, from what I've learned and commented with members of those communities.

Some veteran editors believe that Wikipedia has become more hostile. Do you think that an increase in disputes have led Wikipedia to lose more editors?

I've received extensive feedback on this, as well. Specially, as you say, from veteran Wikipedians feeling that new generations of admins does not share the former spirit of the project. Several threads in Wikimedia foundation mailing list revolve around the same impressions.

I'm waiting until we have a new version of the dumps uncompressed and ready to look into it. In the mean time, I can give you pointers to an interesting quantitative analysis performed by Ed Chi and colleagues at PARC. They presented the paper in last WikiSym 2009 (I chaired that session).

You may also find commentary by Andrew Lih interesting.

You said that the growth of the different language Wikipedia is the same regardless of policies. But you also cite Andrew Lih, who believes that Wikipedia has become too insular and closed. Isn't this a contradiction? Or is there some other explanation of what has caused the community to stabilize?

Yes, there is, let me explain it.

I believe that, indeed, the Wikipedia community is getting less and less friendly and welcoming for new editors. All feedback from editors/admins, messages and evidences of disputes in mailing lists and reverts point to that direction. New admins only "heard of Jimmy Wales' philosophical statements indirectly, but they did not lived the spirit of the project start first-hand" (this is more or less the overall argument behind many discourses).

So, (and this is a delicate issue) my point is that the problem is not in adding or removing policies. The problem is that new people coming does not share the same spirit of veteran admins (e.g. assume good faith, obviate all rules if it is in favor of improving the project, search for consensus and leave aside personal disputes, etc.). They put more difficulties to open new articles, edit the articles under "their own control", and debate with other people.

Therefore, our results just showed us that this is true: policies have no influence on this stabilization effects. Instead, what is changing is the behavioral patterns in the community (take a look at the metrics by PARC folks, showing the raise of reverts; the method is not exact, but the numbers are reasonable proxies, in my view). No contradictions found here.

For sure, I think that this can be mitigated with concrete policies and grass roots support. In particular:

- * Taking better care of newbies. Easier-to-find path to sandboxes, tutorials and basic policies. Mentoring.
- * Specific strategies for contributions from academia (from High School, to Universities to R&D centers).
- * Improving usability and contextual information in MediaWiki (like WikiTrust, WikiDashboard and similar add-ons). Better support for admin tasks.

Finally, I'd also like to point out the need for including explicit support of social network contacts, if not for all users, at least for Wikipedia admins. It would be great if the interface allows you to quickly find out who's writing in a certain article, possibly connected to other people you may know directly.

Many tools for collaborative content creation are starting to care about this issue of social network support.