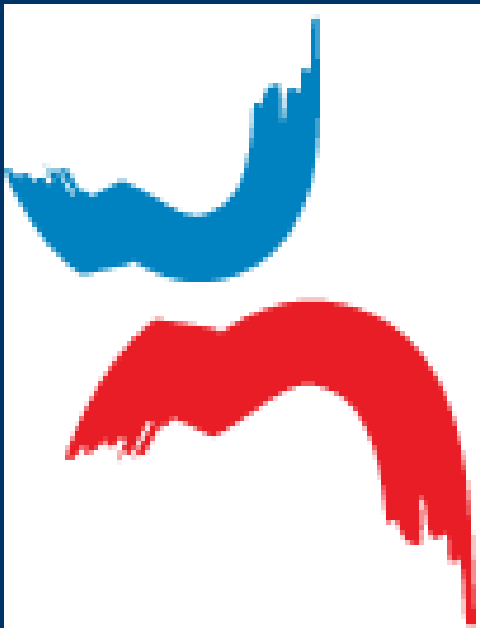


*Where Have All The Writers
Gone:
The Diversion, Distraction, and
Departure of Wiki “Content
Creators”*



Seth Anthony
[[en:User:Sethant]]

Wikimania 2007 Taipei
3 August 2007

Who writes Wikipedia?

"There is this tight community that is actually doing the bulk of all the editing. It's a group of around a thousand to two thousand people." – J. Wales, 2005

“It's the outsiders who provide nearly all of the content [...] then insiders make several edits tweaking and reformatting it.” - Aaron Swartz, 2006

“The really important query is 'Who cares who writes wikipedia?’” - blog comment

“Not all edits are created equal...”

Edits with negative value:

- Vandalism/”trolling”/disruption

Edits with positive value:

- Style/formatting/categorization -- “tweaking”
- Addition of content

Edits with possible value:

- Administration/bureaucracy
 - Discussion/dispute resolution
-
-

“Thousands of editors at thousands of keyboards...”

Classification of a sample of 250 edits:

Outside article namespace	28%
Article talk namespace	10%
Article namespace	62%
Vandalism	5%
“Tweaking”/minor changes	45%
Content addition	12%
Article creation	2%

Who are our content creators?

Of high-content edits:

- 0 were made by admins
- 69% are registered users
- 52% have a userpage (Wikipedia “identity”)

Median date of first edit: ~4 months.


Median number of edits: ~200

Do collective statistics tell the whole story?

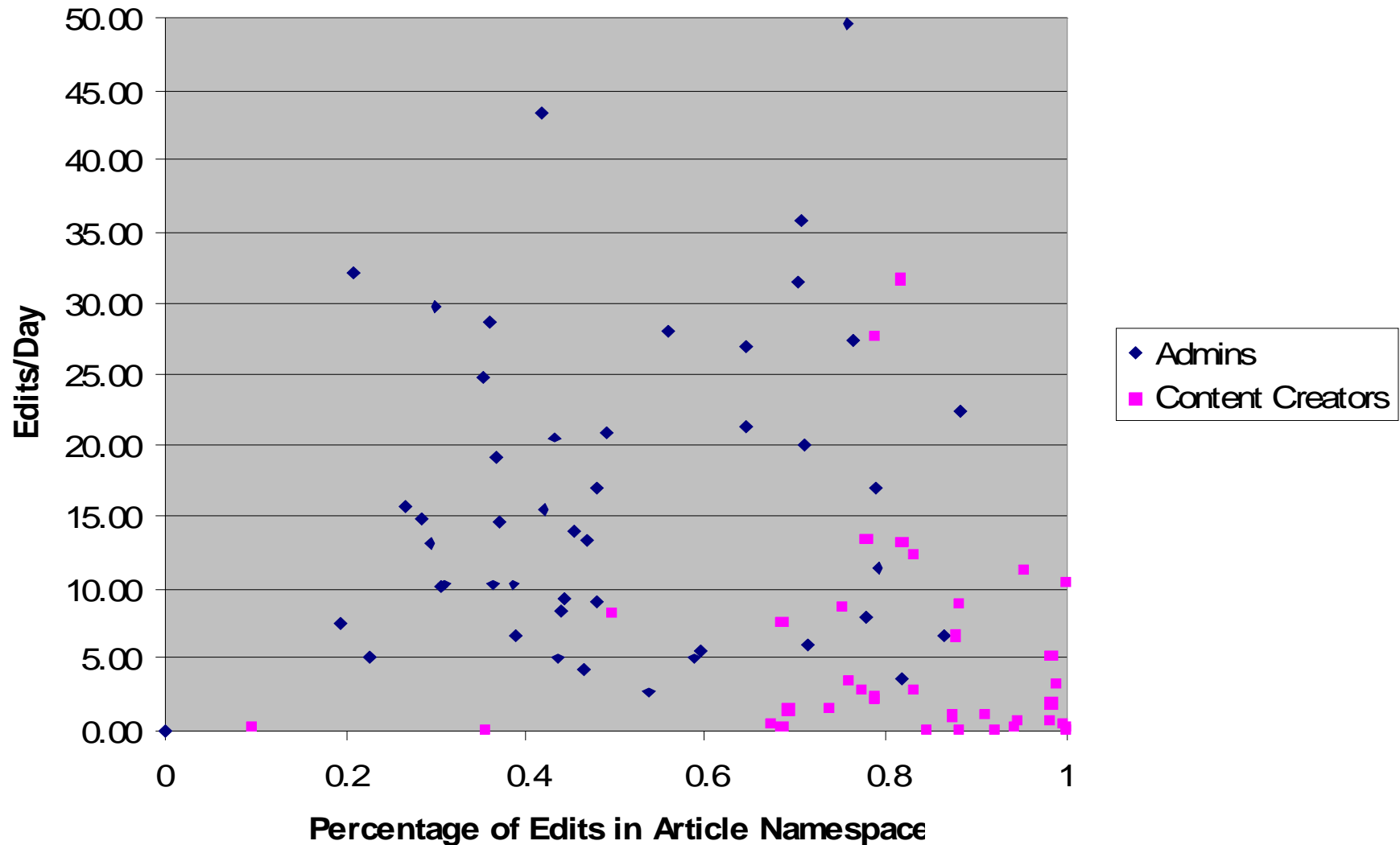
Can we identify – and do we value – all those who add valuable content to our wikis?

“Content creators” v. “admins”

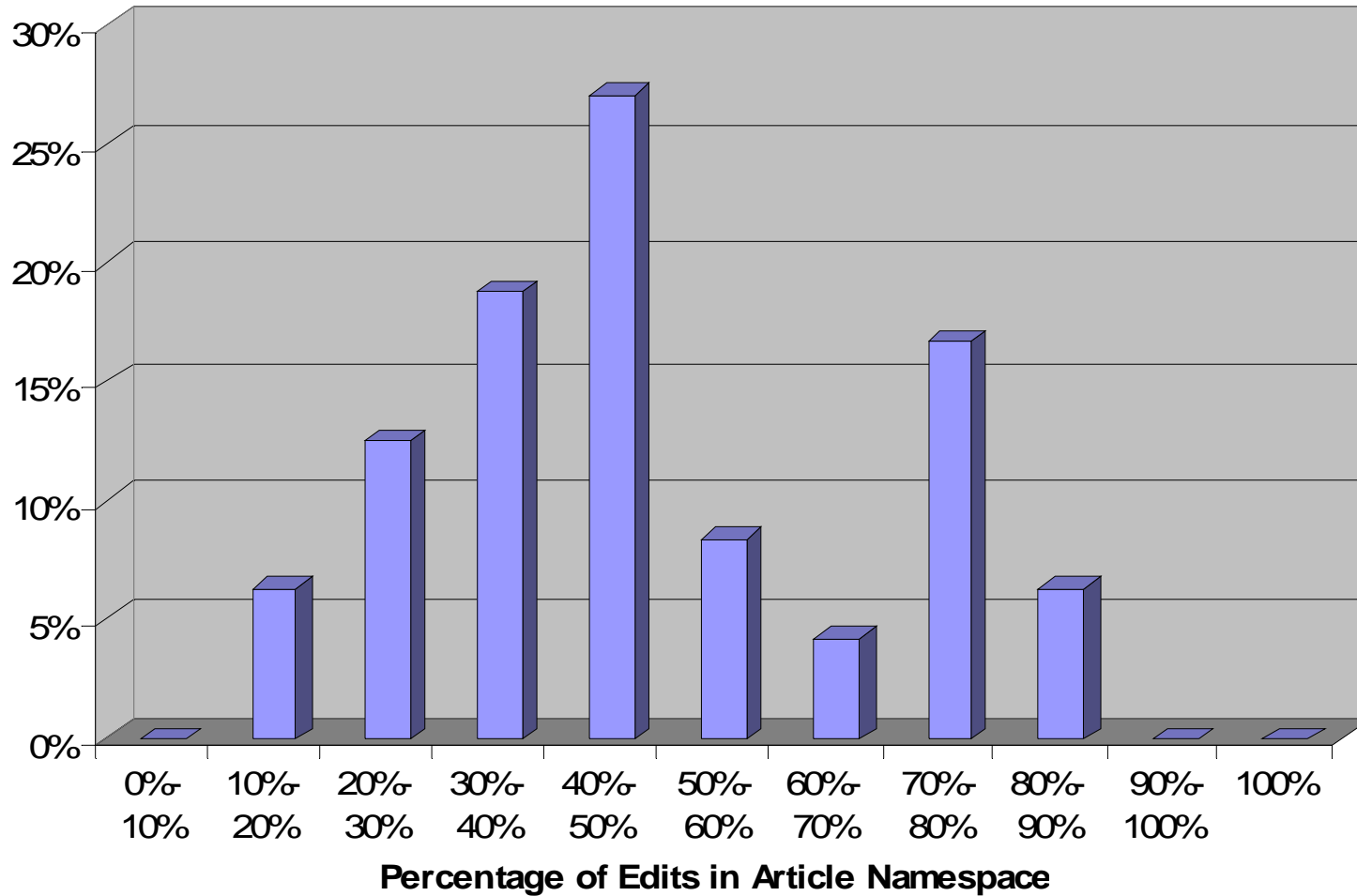
# edits:	1700 (avg.) 105 (median)	12900 (avg.) 11303 (med.)
% mainspace:	81%	51%
edits per page:	2.27	2.20
time editing:	251 d. (avg.) 145 d. (med.)	923 d. 828 d.
edits per day:	5.23	16.6



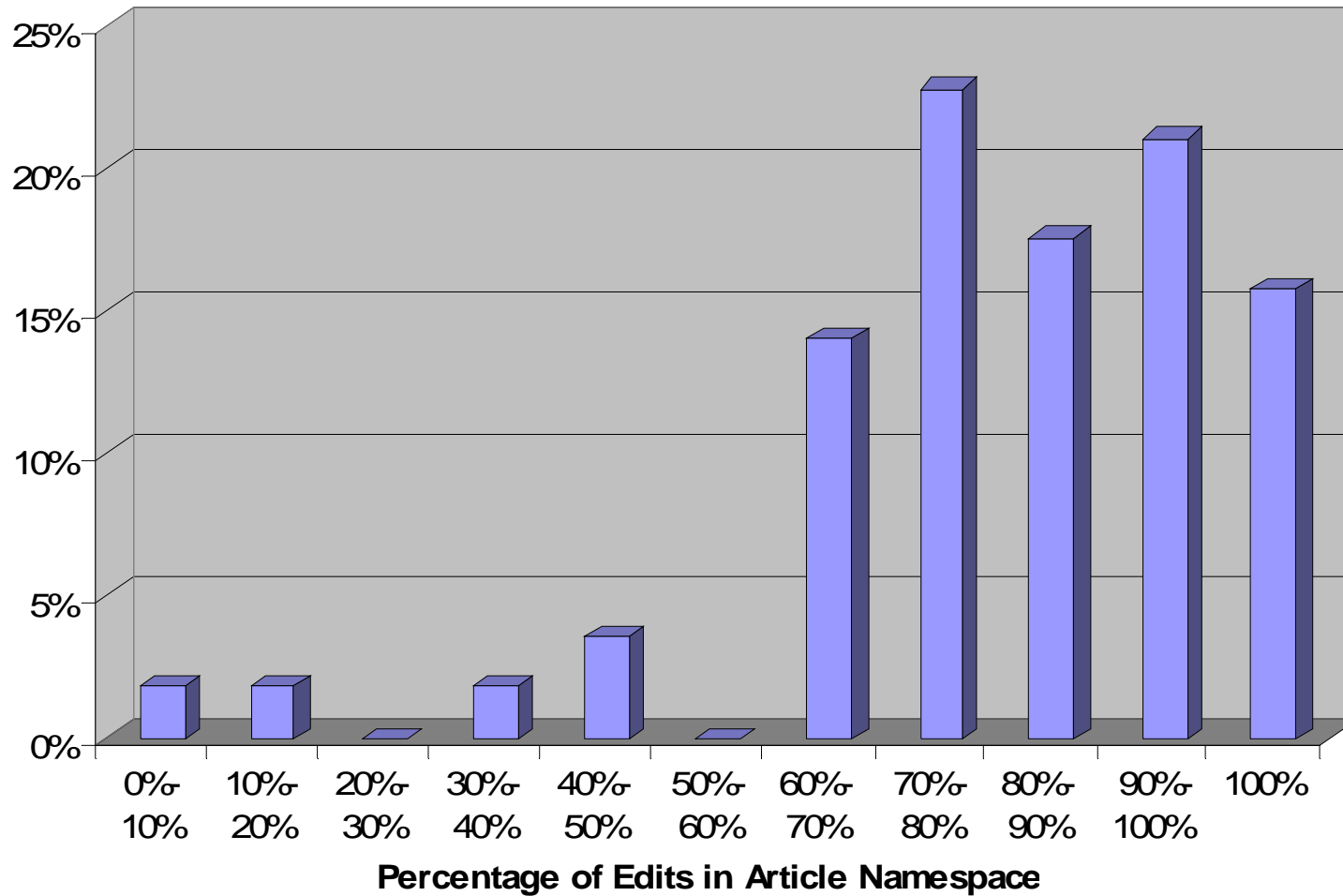
“Content creators” vs. “admins”



Editing Articles: Admins



Editing Articles: Content Creators



Content Creators I: Anons

24% of content edits are from accounts:
fewer than 50 edits
accounts (or use of IP) < 1 week old

Half are first edits.

Half never edit after that day.

Half are anonymous users; others have no userpage.

Are they: One-time editors? Accidentally logged out?

Content Creators II: “Dabblers”

28% of content edits are from accounts:
fewer than 150 edits
editing more than 1 month

Average < 1 edit per day.

~80% have edited within the last month.

Less likely to have a userpage/community identity

Given more free time, they might be...

Content Creators III: Wikipedians

48% of content edits are from accounts:
More than 1 edit/day

Almost all have edited w/in the last week,
or edited actively for > 1 year.

Class includes almost all new article creators.
Editors tend to have a “focus” topic area.

Total Wikipedia career generally includes > 500 edits

Range of editing rates: 1/day to 30/day
and edit counts: 300 to 30,000

Members (or future members) of the “gang of 1000”?

Admins A and B:

Bimodal distribution of admin activity, measured by fraction of edits in article namespace

Admins A: 20%-60% in article space
70% of admins

Admins B: 60%-90% in article space
30% of admins,
similar distribution to “classic editors”

Q: Were they once “Wikipedians” who changed their focus to administrative tasks?

How did they start?

	'Pedians	Admins B	Admins A
Edits/Day	7.8	17.1	14.5

Article Namespace Edits:

Overall:	82%	75%	42%
First 50:	91%	91%	62%
Recent 50:	76%	72%	37%

'Pedians	Admins B	Admins A
----------	----------	----------



Summary: Identifying Editors

Content Creators I:

- anonymous users/drive-by editors

Content Creators II:

- occasional dabblers

Content Creators III/Admins B:

- subject area editors

Admins A:

- professional bureaucrats

Editors (and admins) are not homogenous.

50% of content edits

come from outside the “community”

Future directions and acknowledgements

Data mining (distributed classification of edits?)

Distribution of activity within in classes
(non-article/content/non-content)

Comparative studies:

Wiki size/stage of development

Languages/Types of wikis

Thanks to:

- Wikipedia contributors, readers, and critics
 - Interiot and other tool creators
 - Colorado State University
-
-

Licensing



This file is licensed under the terms of the

GNU Free Documentation License:

<http://www.gnu.org/copyleft/fdl.html>

and

the Creative Commons-BY-SA license:

<http://creativecommons.org/licenses/by-sa/2.5/>

“Non-Content” edits: Another piece of the puzzle

Recent content edits on en.wikipedia
(from July 2007: “Recent changes”):

Same three classes:

Editors I (anons) – 32%

Editors II (dabblers) – 12%

Editors III ('pedians) – 56%